AD NUMBER

ADB238010

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Information; Aug 97. Other requests shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott St., Fort Detrick, MD 21702-5012.

AUTHORITY

USAMRMC ltr dtd 26 Jan 2000

AD_____

AWARD NUMBER DAMD17-94-J-4025

TITLE: The Role of the EGF Receptor First Intron in Its Regulation in Breast Canceer

PRINCIPAL INVESTIGATOR: Susan A. Chrysogelos, Ph.D.
Melissa Ann Wilson

CONTRACTING ORGANIZATION: Georgetown University
Washington, DC 20007

REPORT DATE: August 1997

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1997	Final (1 Sep 94 - 31 Jul 97)

4. TITLE AND SUBTITLE
The Role of the EGF Receptor First Intron in Its Regulation in Breast Cancer

5. FUNDING NUMBERS
DAMD17-94-J-4025

6. AUTHOR(S)
Susan A. Chrysogelos, Ph.D.
Melissa Ann Wilson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Georgetown University
Washington, DC 20057

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT: Distribution authorized to U.S. Government agencies only (proprietary information, Aug 97). Other requests for this document shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, Maryland 21702-5012.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

Hormone independent human breast cancer is characterized by estrogen receptor (ER) loss and the acquisition of high epidermal growth factor receptor (EGFR) levels. Despite the tendency for an inverse correlation between EGFR and ER, EGFR is a strong prognostic indicator for poor survival rate independent of ER status suggesting that EGFR overexpression is an important step in the progression to estrogen independence. Our studies have shown that several DNase I hypersensitive sites which correspond to potential regulatory protein binding sites reside within the EGFR gene first intron exclusively in hormone independent breast cancer cells, and micrococcal nuclease assays indicated that a disrupted and shifted nucleosome phasing pattern of the EGFR first intron occurs in a high EGFR expressing cell line. CAT assays demonstrated that a 140bp region of the first intron of EGFR has enhancer ability specifically in these hormone independent breast cancer cells. The DNA-protein interaction that occurs in this region was localized to a 35bp fragment that retains enhancer activity, and potentially to the sequence ATGACT. We hope that identifying the specific regulatory elements involved in EGFR up-regulation will assist in developing new therapies for preventing and treating aggressive, estrogen independent breast cancer.

14. SUBJECT TERMS
Breast Cancer; EGF Receptor; Transcription; Enhancers; Intron Elements; Estrogen Independence

15. NUMBER OF PAGES
56

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

SAE Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

SAE In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

SAE In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Susan A. Sheppeller 6/1/98

PI - Signature Date

THE ROLE OF FIRST INTRON ELEMENTS IN EGF RECEPTOR REGULATION IN HUMAN BREAST CANCER

TABLE OF CONTENTS

THE ROLE OF FIRST INTRON ELEMENTS IN EGF RECEPTOR REGULATION IN HUMAN BREAST CANCER

INTRODUCTION

Breast cancer often progresses from an estrogen receptor (ER) positive, hormone dependent phenotype that is sensitive to anti-estrogens, to a more aggressive, ER negative, hormone independent phenotype that does not respond to anti-estrogens (1, 2). It is hormone independent breast cancer, no longer under the control of estrogen, that is the focus of most studies since there is no specific target for therapy in these tumors. The loss of ER in the hormone independent state implies that the cell utilizes other pathways to bypass the requirement of estrogen for proliferation. One theory is that the up-regulation of growth factors and/or their receptors in these cells provides a mechanism for escaping estrogen dependence (3, 4). Therapies that will be developed for hormone independent breast cancer will ultimately attempt to suppress growth of ER negative cells by blocking these alternative pathways for cell growth.

The epidermal growth factor receptor (EGFR) is one such receptor that has been well documented to be overexpressed in several types of tumors (5). In human breast cancer, the estrogen independent state is characterized by the loss of ER and the acquisition of high levels of EGFR (1, 2, 6). The EGFR status of a tumor has been shown to be a better predictor for more aggressive tumors, poor rate of survival and failure of endocrine therapy than ER status (7-9). Additionally, tumors that do not show an inverse relationship between ER and EGFR almost exclusively co-express the two receptors (10), strongly supporting the hypothesis that overexpression of EGFR is a critical step in the progression to estrogen independence.

The EGFR gene is 110kb and is located on chromosome 7p 12-14 (11). It has 26 exons and the first intron, which we will be focusing on, is 18kb alone. The GC rich promoter of EGFR does not have a TATA or CAAT box (12). The lack of a specific binding site for the TATA binding protein that initiates the assembly of the transcription machinery results in the occurrence of multiple transcription start sites (11, 12). In breast cancer, EGFR is often overexpressed but rarely amplified at the gene level (13), and total EGFR message levels correlate with receptor binding sites/cell (6). Nuclear run-off data has shown that transcriptional control is an important regulatory mechanism for EGFR expression in these cell lines (6).

The large majority of the studies that have been performed on EGFR transcriptional regulation have been done using HeLa cells, and A431 cells which were derived from an epidermoid carcinoma and have an amplified and rearranged EGFR gene (14). In vitro transcription assays and binding assays using A431 nuclear extracts demonstrated that numerous factors are capable of interacting with the EGFR promoter (15). Most of the transcription factors that have been shown to enhance activity through the EGFR promoter are general transcription factors but their activities can not account for the characteristic expression of EGFR in ER negative breast cancers (15-18). Further experiments using transient transfection assays in the HeLa cells looked at the role of the promoter and first intron in EGFR expression. Two regions, one upstream from the major transcription start site and one downstream in intron 1 showed enhancer activity individually (19). These enhancers acted synergistically when they were in the same construct (19), and each corresponds to DNase I hypersensitive sites in A431 cells, implying that regulatory factors bind to these regions in vivo. We believe that it is necessary to look at the regulation of EGFR in breast cancer cell lines specifically to see what elements are involved in its overexpression in these cells.

Studies in our laboratory have examined the differences in regulation of the EGFR gene in ER positive and ER negative breast cancer cell lines utilizing DNase I hypersensitivity assays. Numerous investigations have shown that DNase I hypersensitive sites are correlated with the binding of regulatory factors (20). While hypersensitive regions were found in the promoter and first exon in all the cell lines tested, only the ER negative cell lines that express high levels of EGFR were found to have hypersensitive sites in the first intron. Our preliminary data in the original proposal showed through transient transfections with CAT reporter constructs that regions of the EGFR first intron have the potential to regulate EGFR expression in both ER negative and ER positive breast cancer cell lines. Sequencing and analysis of the 2.5 kb region of the first intron allowed us to identify useful restriction sites as well as putative regulatory elements for general transcriptional factors that might interact with these intron regions. We have also found through micrococcal nuclease assays a disrupted and shifted nucleosome phasing pattern in the first intron of EGFR in hormone independent cells compared to the canonical phasing observed in hormone dependent cells which concurs with the aforementioned DNase I hypersensitivity data.

Hypothesis

The long range goal of this project is to understand the molecular mechanisms involved in the overexpression of EGFR in hormone independent breast cancer. The overexpression of EGFR in ER negative breast cancer is under transcriptional control and therefore under the control of specific regulatory factors. Our studies have identified potential regulatory regions in the 5' region of the first intron of EGFR in breast cancer cells with high levels of EGFR. We hypothesize that there are specific elements within the first intron of EGFR that are involved in its up-regulation. The objective of this work is to locate and characterize the cis elements in the first intron of EGFR that are involved in its transcriptional regulation in hormone independent breast cancer cells. The identification of these regulatory elements could, in time, lead to modes of therapy involving the targeting of the overexpression of EGFR in hormone independent breast cancer.

Methods of Approach

The experiments described in this report were designed to locate and identify the specific elements in the EGFR first intron that are involved in its up-regulation in hormone independent breast cancer. Our approach involved a combination of in vivo and in vitro techniques, such as micrococcal nuclease assays to determine nucleosome phasing, transient transfection assays with CAT reporter constructs to determine regulatory ability, and DNase I footprinting to identify specific protein binding sites.

MATERIALS AND METHODS

Cell Culture

All breast cancer cell lines were obtained from the Lombardi Cancer Center Tissue Culture Core Facility. The ZR-75-1, BT474, T47D, MDA-MB-231, MDA-MB-468, BT549, and HS578T breast cancer cell lines were originally obtained from the American Type Culture Collection (ATCC, Rockville, MD) as was the COS-1 African green monkey kidney cell line. The MCF-7 cell line was originally obtained from Dr. Marvin Rich (Michigan Cancer Foundation) and the MCF-7/ADR cell line was originally obtained from Dr. Kenneth Cowan (Clinical Pharmacology Branch, NIH, Bethesda, MD). The T47Dco cell line was obtained from Dr. Dean Edwards (University of Colorado, Denver, CO). All cell lines were maintained in Richter's modified minimal essential medium (IMEM; Biofluids, Rockville, MD) supplemented with 10% fetal bovine serum (FBS; Biofluids). The cells were maintained at 37°C in a humidified atmosphere of 95% air-%5 CO_2 and the media was changed approximately every three days. The cells were harvested by trypsinization and split 1:15 approximately every seven days.

Sequencing and Analysis

All of the constructs for sequencing were subcloned from two clones we received from Dr. Glenn Merlino (National Cancer Institute, NIH, Bethesda, MD). One is a chloramphenicol acetyltransferase (CAT) reporter construct, pERCAT2DE (21) which contains the EGFR gene from -2200 to -15 and from +398 to +2577, and the other, pEP1, is a genomic clone from a human fetal liver cell library (12). This genomic clone spans the EGFR sequence from -1000 to +2800. From pEP1, 1100bp, 600bp and 500bp SstI fragments of the 5' region of the first intron were subcloned into the SstI polylinker site of pBluescript II KS (Stratagene, La Jolla, CA). Also, a 730bp PstI fragment and a 500bp SstI-EcoRI fragment from the first intron were subcloned into the PstI and SstI-EcoRI polylinker sites of pBluescript II respectively. Additionally, individual constructs containing the 780bp and the 660bp intron 1 PstI fragments from pEP1 subcloned into the pCAT-Promoter (Promega Co., Madison, WI) PstI site in the enhancer polylinker were used to complete the sequencing. All restriction enzymes were purchased from New England Biolabs Inc. (Beverly, MA).

The plasmid DNA for sequencing was prepared by conventional alkaline lysis, and sequenced by the Sanger dideoxy method with reagents from the Sequenase 2.0 kit (United States Biochemicals, Cleveland, OH). When the simple dideoxy method was unable to adequately resolve the GC rich 5' region of the intron, PCR sequencing was performed with the fmol Sequencing System (Promega Co., Madison, WI) according to the manufacturer's instructions. The pUC/M13 forward primer (5'-GTTTTCCCAGTCACGAC-3'), which came with the Sequenase 2.0 kit, and the pUC/M13 reverse primer (5'-AACAGCTATGACCATG-3') were used to sequence the constructs containing the plasmid backbone of pBluescript II while only the reverse primer was used to sequence the pCAT-Promoter clones. The reverse primer was manufactured by the Lombardi Cancer Center Sequencing and Synthesis Facility. We utilized the University of Wisconsin GCG database to analyze our sequencing results.

Constructs

Throughout this work the schematic diagrams of the regions of the EGFR first intron that are subcloned into our constructs show the DNase I hypersensitive (DH) sites we have found in the first intron of the EGFR gene in hormone independent breast cancer cells (22). Within the boxes that depict the DH sites are light and dark regions that represent fine mapping of the DH sites that was performed with the native genomic blotting technique (23,24). This technique permits the visualization of both DH sites, shown as dark bands, and protein footprints, shown as light protected bands.

Our CAT constructs were subcloned from the two clones mentioned in the previous section, pERCAT2DE (21) and pEP1 (12). From the pERCAT2DE construct, the approximately 2.1kb 5' flanking region and promoter of EGFR (-2200 to -15) was subcloned into the XhoI site of pGEM-7Zf (Promega Co., Madison, WI) and named EGFR-5'Xho. A 2.1kb region of the first intron of EGFR from the pERCAT2DE construct was also subcloned into the BamHI site of pGEM-4Z (Promega Co., Madison, WI) and named EGFR-3'Bam. All restriction enzymes were purchased from New England Biolabs Inc. (Beverly, MA).

SV40 promoter-EGFR intron constructs: This series of CAT constructs were created with the pCAT-Promoter plasmid (Promega Co., Madison, WI) which contains the SV40 promoter proximal to the CAT gene and an enhancer cloning site distal to the CAT gene into which regions of the EGFR first intron were subcloned. For our purposes we have renamed the pCAT-Promoter plasmid pSV-CAT and we use it as our parental vector for this series of constructs. Our pSVI-CAT construct was made by subcloning the 2.1kb BamHI fragment of the first intron of EGFR from the EGFR-3'Bam construct into the BamHI

enhancer polylinker site of pSV-CAT. The $pSVI_1$-CAT construct contains the 730bp 5' intron 1 PstI fragment from the genomic clone pEP1 subcloned into the PstI site in the enhancer polylinker of pSV-CAT. Our $pSVI_2$-CAT and $pSVI_3$-CAT constructs contain the more distal 780bp and 660bp PstI fragments respectively from the EGFR-3'Bam subcloned into the PstI site in the enhancer polylinker of pSV-CAT. The $pSVI_{2+3}$-CAT construct was made by a partial PstI digestion of EGFR-3'Bam and the resulting 1440bp PstI fragment was subcloned into the PstI site in the enhancer polylinker of pSV-CAT. The constructs $pSVI_{3a}$-CAT and $pSVI_{3b}$-CAT were made by subcloning the 500bp XbaI fragment and 140bp XbaI fragment from $pSVI_3$-CAT into the XbaI site of the enhancer polylinker of pSV-CAT respectively. The pSVIΔ3b-CAT construct was made from the pSVI-CAT construct by removing the 140bp XbaI fragment (3b) and religating.

EGFR promoter and intron constructs: This series of CAT constructs originates from pJFCAT which is a CAT expression vector we received from Dr. Judith Fridovich-Keil (Emory University, Atlanta, GA) that has three polyadenylation sites just upstream of the promoter cloning region to insure that there is no read through transcription (25). The EGFR promoter that is used to make this series of constructs is an approximately 840bp fragment that extends to position -15 relative to the start of translation and contains all transcription start sites mapped in vivo and in vitro (11,12). This 840bp BglII-XhoI fragment from the EGFR-5'Xho clone described above was subcloned into the promoter polylinker sites of pJFCAT. This plasmid, pJFE-CAT, is the parental vector for the constructs made with the endogenous EGFR promoter. Constructs $pJFEI_4$-CAT, $pJFEI_5$-CAT and $pJFEI_6$-CAT were made from a SstI digestion of the pEP1 genomic clone, and the 1100bp, 600bp and 500bp SstI fragments respectively were subcloned into the SstI site of the enhancer polylinker of pJFE-CAT.

Transient Transfection

Plasmid DNA was prepared for transfection by a Maxiprep kit (Promega Co., Madison, WI) or a modified alkaline lysis procedure. In this procedure the DNA was extracted by conventional alkaline lysis, passed through a 0.2mm syringe filter, purified by 5M LiCl precipitation, incubated with 10mg/ml RNase, and then precipitated with 1.6M NaCl and 13% polyethylene glycol. This precipitate was extracted with 1:1 phenol/chloroform twice and chloroform once and ethanol precipitated. All DNA for transfection was quantitated by spectrophotometry and by electrophoresis in a 1% agarose gel. The pAd2CAT (26) that was used for the preliminary selection of cell lines was obtained from Dr. Dorraya El-Ashry (Lombardi Cancer Center, Georgetown University, Washington, DC). The plasmid pCAT-control (Promega) was also used as a positive control for transient transfections and transfection efficiency.

Cells were plated for transfection in 100mm dishes at a density of approximately 1 to 5 x 10^6 (depending on the cell line) and were transfected by either the calcium phosphate method (27,28) or the lipofectamine method (29) when the cells reached 80-85% confluency. For each dish, the calcium phosphate method involved incubating 0.5ml of 0.25M $CaCl_2$ containing 20μg of DNA (of which 5μg was DNA of interest and 15μg was carrier DNA [sheared herring testes DNA, Sigma, St. Louis, MO]) with 0.5ml of HEBS buffer (280mM NaCl, 50mM HEPES pH7.0, 1.5mM Na_2HPO_4) for 30 minutes at room temperature. This 1ml solution was then added dropwise to the dish that had been replenished with new media, and was incubated for 18-24 hours. The lipofectamine method involved separately diluting 10μg of DNA in 0.5ml serum free media and 20ml of lipofectamine (Gibco/BRL, Gaithersburg, MD) in 0.5ml serum-free media, and then mixing them together. The mixture was incubated for 30 minutes at room temperature, diluted with 3ml of IMEM supplemented with 5%FBS and then placed on the cells for an incubation period of 18-24 hours.

For both methods, after the transfection incubation, the media was replaced with IMEM supplemented with 10% FBS and 1% penicillin/streptomycin (Biofluids, Rockville, MD) for another 24 hours. The cells were then washed with phosphate buffered saline (PBS; Gibco/BRL, Gaithersburg, MD) and harvested in TNE (0.04M Tris-HCl, pH 7.5, 0.15M NaCl, 1mM EDTA). The cell pellets were divided in half so that one half can be used for the CAT assay and the other half can be used for transfection efficiency measurement. The cell pellets were stored at -70°C.

Transfection efficiency: Throughout this work, transfection efficiency was measured by the Hirt method (30), by the β-galactosidase enzyme assay system (Promega Co., Madison, WI) and by our own transfection efficiency method (31). The Hirt method involved lysis of the cells in a lysis buffer (0.01 Tris-HCl, pH 7.5, 0.01M EDTA, 0.6% SDS) and incubation of the lysed cells with 1M NaCl on ice overnight. After centrifugation the supernatant was extracted with 1:1 phenol/chloroform and ethanol precipitated. The plasmid DNA was then used to perform a Southern blot as described below. The β-galactosidase enzyme assay system is a colorimetric assay that was performed according to the manufacturer's instructions. In order to do this assay the plasmid pCHCβ-gal (a gift from Dr. Francis Kern, Lombardi Cancer Center, Georgetown University, Washington, DC), was co-transfected at a concentration of 1μg per dish of cells.

Our method of measuring transfection efficiency is methodologically similar to the Hirt assay except that total DNA, rather than just plasmid DNA, was extracted from the transfected cells. The histone H4 gene, an endogenous cellular gene that has several copies per cell, was used as an internal control. The cell pellets were resuspended in 0.5ml lysis buffer (100mM NaCl, 20mM Tris-HCl, pH 8.0, 20mM EDTA, 0.2% SDS, 0.2 mg/ml Proteinase K [Boehringer Mannheim]) and ocillated at room temperature for five hours. This mixture was then extracted with 1:1 phenol/chloroform twice, chloroform (EM Science) once and ethanol precipitated overnight. The pellet was resuspended in 0.1ml TE (10mM Tris-HCl, pH 8.0, 1mM EDTA) with 200μg/ml RNase A and incubated at 37°C for 5 hours. The DNA mixture was diluted with 0.1ml TE, extracted and precipitated as before. The resulting DNA pellet was resuspended in 30-50ml of TE, and half of the resuspended DNA was digested with EcoRI. The digested DNA was electrophoresed in a 1% agarose gel, and a Southern blot was performed as described below. The blot was probed with a random primed fragment for the endogenous H4 gene and exposed to film. The blot was then stripped and reprobed with diluted, nick translated probe made from the plasmid for the CAT reporter construct and exposed to film. Densitometric analysis of the bands was performed to calculate the ratio of plasmid to cellular DNA, and this ratio was then used to normalize the transcriptional activity data from the reporter gene.

Southern Blots

Southern blotting was performed as detailed previously (22). Purified and restriction digested DNA was electrophoresed in a 1-1.5% agarose gel and alkaline tranferred to nylon (Zetabind: Cuno, Meriden, CN). Probes were labeled either by random oligonucleotide priming or by nick translation. Hybridization was carried out for 20-24 hours in 5XSSPE (1XSSPE: 180mM NaCl, 10mM $NaPO_4$, pH 8.3, 1mM EDTA), 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.01% BSA, 0.2% SDS, 100μg/ml herring testes DNA, 10% Dextran Sulfate. After hybridization, filters were washed three times for 30 minutes at room temperature with 5mM $NaPO_4$, pH 7.0, 1mM EDTA, 0.2% SDS. They were then washed once for 30 minutes at 60°C with 0.1X SSC (1XSSC: 150mM NaCl, 15mM sodium citrate, pH 7.3), 0.1% SDS, and once for 30 minutes at 65°C with 0.1X SSC, 0.1% SDS. The blots were then exposed to Kodak XAR-5 film for autoradiography. If the blots needed to be stripped of the previous probe they were washed twice in 0.1X SSC, 0.1% SDS at 80°C for one hour.

Chloramphenicol Acetyltransferase Assay

The chloramphenicol acetyltransferase (CAT) assay was performed by the liquid scintillation counting method (32) or by the thin layer chromatography method (TLC) (33,34). For both methods the transfected cells were harvested and whole cell lysates were prepared in 0.25M Tris-HCl, pH 7.8 by freeze/thaw. Once the supernatant has been retrieved the lysates were measured for protein concentration by the Bradford assay (BioRad, Melville, NY) The lysates were normalized for protein concentration and are added to CAT assay reaction mix.

The liquid scintillation counting method was the CAT assay method used for the preliminary selection of cell lines. In this method 0.4μCi of ^{3}H-acetyl CoA was added to a mixture of lysate and 0.25mg/ml chloramphenicol in a scintillation vial and then layered with Econofluor II (Dupont, Boston, Mass). The samples were then counted at room temperature by the scintillation counter at timed intervals for two or more hours. The acetylated chloramphenicol diffuses into the organic fluor layer while the labeled acetyl CoA remains in the aqueous layer undetectable by the scintillation counter. The rate of the enzymatic reaction, which is determined by the increase in counts per minute over time, is directly proportional to the concentration of CAT in the reaction.

For the TLC method, the lysate was mixed with 0.5M Tris-HCl, pH 7.8, 0.125 μCi ^{14}C-Chloramphenicol and 0.5mM acetyl CoA and incubated at 37°C for 2 hours. The reaction mixture was then extracted with ethyl acetate (Fisher, Pittsburgh, PA) and dried. The extracted organic compounds were resuspended in ethyl acetate, spotted on a thin layer chromatography plate (Fisher) and incubated in a 95:5 chloroform/methanol tank. After the solvent had run approximately 3cm from the top, the TLC was dried and exposed to film. The substrate and product spots were then cut out of the plates and scintillation counted. A percent substrate conversion was calculated for normalization to transfection efficiency.

Electrophoretic Mobility Shift Assay

The electrophoretic mobility shift assay or gel shift assay (35,36) involved the incubation of the breast cancer cell line nuclear extracts with short radiolabeled DNA fragments of the EGFR first intron. Nuclear extracts from the breast cancer cell lines were made as described by Dignam et al. (37). HeLa nuclear extract was a kind gift from Dr. Anna Riegel (Lombardi Cancer Center and Department of Pharmacology, Georgetown University, Washington, DC). Nuclei were isolated by mechanical lysis in 0.5% NP-40 and the extracts were made by low speed centrifugation in the presence of proteolysis inhibitors. Binding reactions included 1-10μg of nuclear extract, 1μg E. coli DNA (Sigma, St. Louis, MO) as a non-specific competitor, binding buffer (10mM Tris, pH 7.5, 0.1M KCl, 1mM DTT, 1mM EDTA, 6% glycerol and the protease inhibitors 10mg/ml leupeptin, 3.5mg/ml Pefabloc [Boehringer Mannheim], and 20mg/ml Pepstatin) and 160mM NaCl. The reactions were incubated for 20 minutes at room temperature before 1μl of probe (0.3-1ng) was added to the reaction and incubated for another 10 minutes at room temperature. The probe was labeled by polynucleotide kinase using [γ^{32}P]ATP (NEB, Beverly, MA). The binding reactions were then electrophoresed in a 6% non denaturing polyacrylamide gel. DNA-protein complexes were detected on film by their retarded migration rate, or upward shift relative to the free probe.

Competition experiments were performed with the unlabeled fragments corresponding to the 35mer and to other transcription factor binding sites such as NF-1, SP1 and AP-2 to determine the sequence specificity of the protein binding. The DNA used for competition was added to the preliminary binding reaction. The 35mer and the mutant 35mer oligonucleotides were manufactured by the Lombardi Cancer Center Synthesis and Sequencing Facility. The 35mer consists of the sequence 5'-CTTAGAGGTTATGACTGCCAAGACACCATTTCATG-3'. The mutant 35mer consists of the

sequence 5'-CTTAGAGGTTCATGTAGCCAAGACACCATTTCATG-3' with the underlined portion denoting the scrambled bases of the cis element. The other oligonucleotides were purchased from Stratagene (La Jolla, CA) and have the following sequences: NF-1, 5'-ATTTGGCTTGAAGCCAATATG-3'; SP1, 5'-GATCGATCGGGGCGGGGCGATC-3'; AP-2, 5'-GATCGAACTGACCGCCCGCGGCCCGT-3'. The underlined portions of these oligonucleotides denote the consensus transcription factor binding sequences.

DNase I Protection Assay

The in vitro DNase I protection technique (38,39) involved the radioactive labeling of the 140bp XbaI-PvuII fragment and the 35bp DdeI-HinfI fragment with [α^{32}P]ATP by using the Klenow fragment of DNA polymerase I. Labeling at the 5' terminus was carried out by digesting with restriction enzyme Rsa I to produce 5' end labeling for the 140bp XbaI-PvuII fragment and by digesting the 91bp DdeI fragment with HinfI to produce the 5' end labeling for the 35bp DdeI-HinfI fragment. The labeled fragments (1-5ng) were incubated with 15-60μg of crude nuclear extract from MCF-7 and BT549 cell lines in the presence of 1-3μg of E. coli DNA as nonspecific competitor. This was performed under the same conditions used for the gel shift assays except that the incubation buffer did not contain EDTA. The DNA-protein mixture was then partially digested with 50μg/ml DNase I (Sigma, St. Louis, MO) for two minutes in a 25°C water bath. The reaction was stopped with 5ml of 0.5M EDTA and then the samples were electrophoresed in a 6% non-denaturing polyacrylamide gel. The bands corresponding to free probe and to bound probe were cut out of the gel and electroeluted. The precipitated electroeluted samples were counted for β-emission and aliquotted so that each sample had the same amount of counts per minute. The samples were dried down, resuspended in formamide, and then analyzed on an 8% denaturing sequencing gel alongside a sequencing reaction with the same fragment. The resulting gel was visualized by autoradiography. The regions of altered DNase I digestion (as compared to digestion of the free probe) represent regions of DNA-protein interaction.

Micrococcal Nuclease Assay

The micrococcal nuclease assay was adapted from the method used by Moreno et al. (40). Nuclei were isolated from MCF-7 and BT549 cells by resuspending the cells in reticulocyte standard buffer (RSB: 10mM Tris-HCl, pH 7.4, 10mM NaCl, 3mM $MgCl_2$) containing 0.5% Nonidet P-40. Nuclei were pelleted by centrifugation, washed twice in RSB and resuspended in RSB at approximately $3x10^7$ nuclei/ml. Aliquots of nuclei were incubated with 0.001-0.1U of micrococcal nuclease (United States Biochemicals, Cleveland, OH) in the presence of 1mM $CaCl_2$ at 37°C for 10 minutes. A stop solution (62.5mM EDTA, 2.5% SDS and 2.5μg/ml proteinase K) was added to the reactions and the samples were incubated at 37°C for 5 hours. The DNA was purified by phenol chloroform extraction, treated with 200μg/ml RNase for 5 hours at 37°C and extracted again. The DNA was digested with HindIII and electrophoresed in a 1.5% agarose gel with 15μg of DNA per lane. The gel was Southern blotted as described above and hybridized with a random primed probe corresponding to the 3' end of the intron region of interest. The probed blot was visualized by autoradiography.

RESULTS

Sequencing Analysis

Since we are interested in defining the specific cis elements involved in the regulation of EGFR, we sequenced the upstream 2.8kb portion of the first intron of the EGFR gene (figure 1). The EGFR DNA that was used to sequence this portion of the first intron came from human fetal liver cells. We subcloned overlapping portions of the intron in order to sequence the entire region. Analyzing this sequence with the GCG program enabled us to determine available restriction sites that were used for subsequent subcloning

and to determine putative transcription factor binding sites. Several consensus regulatory elements are within this sequence and figure 2 shows the locations of some of the more common ones. Sp1 is a transcription factor that binds to GC rich regions and is known to activate transcription proximally and distally (41). AP-1 is the cis element that is the binding site for Fos and Jun (42), and AP-2 is a common element that mediates activation by cAMP and TPA (43). NF-1 is a transcription factor that binds to CCAAT elements (44). Interestingly, some of these elements map to hypersensitive regions which makes them good candidates for general transcription factors involved in the expression of EGFR. However, because these are ubiquitous factors we expect that there are other specific factors found in ER negative breast cancer cells that are responsible for the up-regulation of EGFR in these cells.

We also used the GCG database to compare our sequence with the only other previously published sequence (45) of the first intron of the EGFR gene which was from DNA isolated from A431 cells, a vulval carcinoma cell line. The A431 cells have an amplified and rearranged EGFR gene (14). We were interested in seeing if there are sequence differences between the normal EGFR gene and the gene from cancerous cells. In our analysis we found that there were 48 gaps in the sequences collectively and there were 25 base changes between the sequences, but the sequences have a 99.02% similarity. A similar comparison was performed between the promoter sequences of EGFR from a genomic clone of A431 DNA (11) and a genomic clone of human fetal liver DNA (12) and there were 16 sequence differences spanning over 400 bases (11). Additionally, the fidelity of the DNA polymerase utilized for our sequencing is $6x10^{-8}$ (United States Biochemicals, Cleveland, OH). Both of these points indicate that the differences that we found between the sequences of the first intron of EGFR in A431 and human fetal liver DNA were not artifactual and that minute differences may occur in genes that encode for the same protein.

The differences that were found between the sequences of the first intron of EGFR in these different cells account for a change in twelve consensus transcription factor sites. Interestingly, one is a putative c-myc site that is about 2.1kb downstream of the exon 1/intron 1 boundary and another is a putative c-fos/SRE half-site located about 1.7kb downstream of the same boundary (see figure 2). Both consensus elements exist in the normal sequence but not in the sequence from the malignant cells. The discovery of the c-myc site in the regulatory region of the EGFR gene is of particular interest due to the synergy of c-myc and TGFα, a ligand of EGFR, to malignantly transform mammary glands in transgenic mice (46). Future work could determine if these putative oncogene factor binding sites and the mutations that occur in the malignant cell are important for the regulation of the overexpression of EGFR.

Cell Lines

Before we started our investigation of the regulatory ability of the upstream first intron regions of the EGFR gene, we screened a series of ten breast cancer cell lines to determine what cell lines would be optimal for this study. The cell lines MCF-7, T47D, MDA-MB-231, and MDA-MB-468 were derived from pleural effusions of breast adenocarcinomas. The ZR-75-1 cell line was derived from the ascitic effusion of an infiltrating ductal carcinoma. The BT474, BT549 and HS578T cell lines were derived from invasive ductal breast carcinomas. The MCF-7/ADR cell line is a derivative of the MCF7 cell line that was selected for resistance to the chemotherapeutic agent adriamycin. The T47Dco cell line is a variant of the T47D cell line that has been passaged for over 50 generations. The ten cell lines were chosen based on their ER and EGFR status. The ER positive, low EGFR expressors are the MCF-7, BT474, T47D, and ZR-75-1 cell lines (6,47). The ER negative, high EGFR expressors are the BT549, HS578T, MCF-7/ADR, MDA-MB-231, and MDA-MB-468 cell lines (6,48,49). The MDA-MB-468 cell line has an amplification of the EGFR gene (49) which is a relatively rare occurrence in breast cancer. The T47Dco cell line has an anomalous ER that makes the cells phenotypically ER negative, and its EGFR levels are higher than

those observed in the T47D cells (50; Yarden, RI, personal communication).

Screening of these cell lines was initially carried out via transient transfection of the pAd2CAT reporter construct by the calcium phosphate method. The pAd2CAT construct, which has a constitutively active adenovirus major late promoter driving CAT gene expression, was used as a positive control plasmid for transfection. Liquid scinitillation counting CAT assays and Hirt assays were performed to identify the cell lines with optimal CAT activity and DNA uptake. CAT activity was measured in three independent transfections of the cell lines with at least two different preparations of DNA, and statistical calculations were performed with the Sigmaplot software program.

Two cell lines, the BT549 and HS578T cell lines, stood above the rest in their ability to express the CAT gene. The BT549 cell line had a CAT activity of $2.3x10^5$ dpm/mg of cell lysate, while the HS578T cell line had a CAT activity of $2.15x10^5$ dpm/mg. However, Hirt assays showed that approximately 10pg (or less) of DNA was transfected into approximately $5x10^5$ cells. The other ER negative, high EGFR expressors had substantially lower CAT activity ($<3x10^3$ dpm/mg) and also seemed to have less than 10pg of DNA in the cells, except for the MDA-MB-231 cell line which had over 1ng of pAd2CAT DNA. An explanation for this lack of a correlation between plasmid uptake and CAT expression may be the inability of the MDA-MB-231 cell line to manufacture a stable CAT mRNA or protein. We chose to use the BT549 and HS578T cell lines to represent the ER negative, high EGFR expressors due to their ability to express CAT successfully.

The MCF-7 cell line had the next highest CAT activity at $1.6x10^4$ dpm/mg. The other ER positive, low EGFR expressors such as BT474, ZR-75-1, and T47D all had very low CAT activity ($<1x10^3$ dpm/mg). Hirt assays showed that approximately 1ng of DNA was transfected into these cell lines except for the ZR-75-1 cells which had much less DNA. The ER negative, high EGFR expressing T47Dco cell line had results that were comparable to the ER positive, low EGFR expressors. Interestingly, the cell lines that contained a large amount of transfected plasmid were not necessarily able to adequately express the CAT gene as observed with the MDA-MB-231 cells. This observation attests to the competence of the individual breast cancer cell lines to be efficiently transfected and express the CAT gene, but our primary concern was that the cells could measurably express the CAT gene. We chose to use the MCF-7 cell line as our representative ER positive, low EGFR expressor. The ZR-75-1 cells had the next best CAT activity but they have a very slow growth rate so we chose to use the T47D cell line as a second ER positive cell line.

Transient Transfection Assays

These studies began with the use of the calcium phosphate method to transiently transfect reporter constructs due to the familiarity of this method. The pCAT-Control plasmid (Promega, Madison, WI) that utilizes both the SV40 promoter and enhancer to drive CAT gene expression was transfected into the breast cancer cells at concentrations of 1, 2, 4, 5, and 7.5μg. Both the 5 and 7.5μg of pCAT showed high amounts of CAT activity, therefore, 5μg of DNA of interest and 15μg of carrier DNA were used for transient tranfection with calcium phosphate in the chosen breast cancer cell lines. Since the pAd2CAT vector was the original vector used to test the ability of the cells to exhibit CAT activity, the MCF-7/ADR and MDA-MB-231 cell lines were also transfected via the calcium phosphate method with the pCAT-Control plasmid to see if these cell lines could be used to represent the ER negative, high EGFR expressors since they are well characterized breast cancer cell lines. However, these cell lines were unable to produce significant CAT activity even with the pCAT-Control plasmid.

With the introduction of the lipofectamine method of transfection (29), we compared it to the calcium phospate method in both COS-1 and HS578T cells. Transient transfection of 10μg of the pCAT-Control plasmid into these cell lines was performed by both methods. The CAT activity of the reporter construct varied in the HS578T cells from an average of 4% conversion with the calcium phospate method to an average of 43% conversion with the lipofectamine method, and in the COS-1 cells from an average of 8% conversion with the calcium phospate method to an average of 86% conversion with the lipofectamine method. Based on these results, we decided to use the lipofectamine method as our method of transient transfection.

The pCAT-Control plasmid was transfected into the breast cancer cell lines at the concentrations of 1, 2.5, 4, 5, 7.5, 10μg by the lipofectamine method of transfection. The concentration of 10μg of pCAT-Control produced the best results, so all remaining transfections with the lipofectamine method were performed with 10μg of DNA. Since the lipofectamine method resulted in much higher levels of CAT activity than the calcium phosphate method, it was used to transfect the MCF-7/ADR and MDA-MB-231 cell lines with the pCAT-Control plasmid for the same reasons mentioned above. Again, these cell lines were unable to produce significant CAT activity.

Transfection efficiency: When a transient transfection is performed on cells, the transfection efficiency of the plasmid into the cells is typically measured and used to normalize transcriptional activity. We attempted using the traditional method for measuring transfection efficiency, the Hirt assay, which involves selective isolation of the plasmid DNA from the transfected cells (30). However, there is no direct internal control for the number of cells utilized in the assay; as an indirect control for the cell number per sample, cells are counted prior to performing the assay. This is both time consuming and requires the assumption that each isolation is quantitative and reproducible. We also attempted cotransfecting the β-galactosidase gene, another reporter gene, into the cells and measuring its activity as a method of measuring transfection efficiency. We and others have found the β-galactosidase method for determining transfection efficiency to be unreliable for several reasons. β-galactosidase has variable sensitivity depending on isolation and storage conditions (51), and there can also be alterations in β-galactosidase activity when it is cotransfected with a reporter construct presumably due to promoter cross-competition (52). The possibility of promoter cross-competition becomes a problem when dealing with any method of measuring transfection efficiency by cotransfection since different promoters can use the same general transcription factors, and the use or overuse of one promoter could monopolize the abundance of these transcription factors within a cell.

The ideal way to measure the transfection efficiency of a reporter construct into a cell is to detect the transfected DNA itself and to normalize it to an internal control that represents the number of cells that are used for this detection. We have devised a way to accurately determine transfection efficiency utilizing a portion of the transfected cells in an assay that is methodologically similar to the Hirt assay except that total DNA, rather than plasmid DNA, is extracted from the transfected cells (31). A Southern blot of this DNA is probed for both the transfected plasmid DNA and an endogenous histone H4 gene that is present at approximately six copies per cell as an internal control. Densitometric analysis of the bands is performed to calculate the ratio of plasmid to cellular DNA, and this ratio is then used to normalize the transcriptional activity data from the reporter gene.

We have performed this assay with several different cell lines and with different methods of transfection (31). The cell lines that have been used are HS578T, MCF-7, T47D and BT549 human breast cancer cells, and COS-1 African green monkey kidney cells. We have used both the calcium phosphate

and lipofectamine methods of transfection and have found this assay to be satisfactory for measuring transfection efficiency with both methods. To investigate whether our total DNA preparations reflect only transfected plasmid, that is DNA taken up by the cell, or if additional plasmid attached to the exterior of the cells might be copurified, we compared total and nuclear DNA preparations for both calcium phosphate and lipofectamine transfections. We observed no significant difference in the amount of DNA in either the total or nuclear DNA preparations and, therefore, conclude that the DNA extracted with this method is contained within the interior and most probably the nucleus of the cell. We have also tested this method of purification for its ability to extract genomic and plasmid DNA equally by performing total DNA preparations on aliquots of cells that had 1pg, 10pg, 100pg, 1ng, 10ng, 100ng and 1μg of plasmid DNA added to the aliqouts before total DNA preparation. We found that plasmid DNA recovery is proportionally correct over this range and that genomic DNA recovery is reproducible. This demonstrates that the assay does not selectively extract and purify one species of DNA over the other but rather extracts both species of DNA reproducibly and proportionately (31).

This method of measuring transfection efficiency has been most useful in our studies since we are transfecting a series of constructs into different cell lines. We have performed this method of measuring transfection efficiency for the cell lines we used several times and have found that it did not change the overall trend of the data. The CAT activity that is observed with the various constructs is not due to increased transfection efficiency of the different cell lines, but rather to the transcriptional activity of the construct. We found that the cell lines have different transfection efficiencies and different abilities to express the transfected CAT gene, but this does not make it impossible to compare the resulting CAT activity. Using the parental vector as a measure of basal activity within the specific cell and calculating fold induction based on that basal activity allows us to compare the transcriptional activity of the constructs in the different cell lines. Once we had established that normalization did not affect the data produced by our cell lines, we stopped using a normalization technique and instead relied on the data from several CAT assays that were performed with at least three different DNA preparations.

Chloramphenicol Acetyltransferase Assays

The chloramphenicol acetyltransferase (CAT) assay is used to quantitatively measure the ability of a region of DNA to drive expression of the bacterial CAT gene. We have designed several different CAT gene constructs to aid us in delineating the transcriptional activity of the region of the first intron of EGFR that contains the DNase I hypersensitive sites. Our constructs contain portions of the first intron cloned into the enhancer site of a parental CAT expression vector and either a heterologous SV40 promoter or the homologous EGFR promoter. The use of the different promoters allows us to examine which intronic elements are capable of affecting expression with a general promoter and which elements are able to affect expression in the presence of the natural EGFR promoter. The CAT constructs described below were transiently transfected into the breast cancer cell lines expressing high and low levels of EGFR as described in Materials and Methods. The following data is from CAT assays that were performed by the TLC method, and is represented as a fold induction with the activity of the parental vector defined as one. The CAT assay data for each figure was calculated from at least five different CAT assays that were transfected with at least three different preparations of DNA. The Sigmaplot software program was utilized to calculate the standard error of the data and to calculate the p value of selected data via the students t-test. This assay provides quantitative data and enables us to measure the ability of cis elements to regulate transcription.

SV40 promoter and EGFR intron constructs: Figure 3 shows the CAT activity of the pSV-CAT series of constructs in the T47D, MCF-7, HS578T and BT549 breast cancer cell lines. The ER positive,

low EGFR expressing cell lines had an overall lower fold induction of CAT activity than the ER negative, high EGFR expressing cell lines. We have also made constructs that are in the reverse orientation to see if the ability to regulate CAT activity changes as compared to that of the forward orientation. In our assays, when a region was able to induce CAT activity more than 5 fold in the forward orientation, generally the reverse orientation was able to induce CAT activity but not to the same extent. One of the definitions for a classical enhancer is a cis element that is able to induce transcriptional activity regardless of orientation. However, it has been shown that enhancers within introns are often orientation dependent (21,53,54).

A comparison of the CAT activity induced by the first intron regions in the forward orientation in the low EGFR expressing and the high EGFR expressing breast cancer cell lines can be seen in figure 4. The 2.1kb first intron region of EGFR, or construct pSVI-CAT, was capable of inducing expression 4 fold in the MCF-7 cells, 10 fold in the HS578T cells and 7.6 fold in the BT549 cells over the basal activity of pSV-CAT. In the T47D cells, pSVI-CAT only induced 1.3 fold over pSV-CAT. The three regions that comprise the intron had similar trends of activity in the cell lines. Region 1 had a slight repressive activity in BT549 cells while in T47D cells it had the same activity as the parental vector. With the other cell lines it was capable of a minimal induction of 2.1 fold in the HS578T cells and 2.4 fold in the MCF-7 cells. Region 2 induced activity to a slightly greater extent: 1.9 fold in the T47D cells, 2.9 fold in the BT549 cells and 5.1 fold in the HS578T cells. In the MCF-7 cells, region 2 had similar activity to region 1 with a 2.4 fold increase in CAT activity over the parental vector. It can be seen that region 3 had the highest amount of enhancer activity of the regions of the intron in all of the cell lines. Region 3 induced activity 2.5 fold in T47D cells, 3.9 fold in MCF-7 cells, 5.5 fold in BT549 cells and 8.6 fold in HS578T. Region 2 seemed to have a slightly repressive effect on the activity of region 3 in all of the cell lines as observed with the construct $pSVI_{2+3}$-CAT. These results demonstrated the presence of enhancer activity within the intron region 3 of the EGFR gene especially in the cell lines that express EGFR at high levels.

In figure 5 we see that further division of region 3 allowed us to delineate a 140 bp region that accounts for the enhancer activity of region 3 in the BT549 cells. Region 3b was capable of inducing CAT activity 12.6 fold while region 3a showed no significant induction in the BT549 cells. Meanwhile, region 3b was only capable of inducing CAT activity 3.5 fold in the MCF-7 cells as compared to a 2 fold induction seen with region 3a. Additionally, there was a significant difference in the CAT activity detected between the cell lines with region 3 ($p=0.013$) and with region 3b ($p=0.019$). When region 3b was removed from the 2.1kb intron region, the ability of the remaining intron region to induce CAT activity decreased from 7.7 fold to 1.7 fold in the BT549 cells (figure 6). This is a significant reduction as measured by Sigmaplot's students t-test ($p=0.05$). In the MCF-7 cells, the loss of the region 3b from the 2.1kb intron did not cause a significant decrease in CAT activity. This DNA fragment, region 3b, that has enhancer-like activity in the high EGFR expressing BT549 cell line was used as a probe in the following gel shift assays and in vitro DNase I protection assays.

EGFR promoter and intron constructs: The parental construct, pJFE-CAT, that contains the EGFR promoter driving the expression of the CAT gene had high levels of activity in the T47D, MCF-7 and BT549 cells. This indicates that cell specific regulation does not occur through this proximal promoter of EGFR. In the T47D cell line, this series of constructs containing intron 1 regions did not have more than 2 fold activity, with region 4 having some inductive ability over the promoter alone in the forward orientation and region 6 showing slight inductive abilities in both the forward and reverse orientations. However, with both the MCF-7 and BT549 cells region 6 showed greater induction in the forward orientation than the reverse (figure 7). In the MCF-7 and BT549 cells, regions 4, 5 and the combination

of 5 and 6 did not have transcriptional activity over the high basal levels of the promoter (figure 8). Region 6, however, was able to induce activity 1.9 fold in the T47D cells, 3 fold in the MCF-7 cells and 2.3 fold in the BT549 cells. Interestingly, region 5 was capable of reducing the activity of both the promoter alone and the promoter in combination with region 6. This series of constructs does not contain region 3b (constructs with region 3b will be described below), and indicates a lack of significant enhancer activity in the more 5' portion of the intron when in conjunction with the EGFR promoter.

Electrophoretic Mobility Shift Assay

The gel mobility shift assay (35,36) was utilized to investigate the ability of the regions of the first intron of EGFR that showed enhancer activity in CAT assays to bind protein. Radioactively labeled fragments corresponding to these regions were incubated with nuclear extracts from the breast cancer cell lines, and DNA-protein complexes were detected by their retarded migration rate, or upward shift relative to the free probe. This assay allowed us to define regions of the first intron of EGFR that can bind nuclear factors and to attempt to identify factors that are specifically found in the cell lines expressing high levels of EGFR.

Since regions 2 and 3 demonstrated the ability to upregulate CAT activity, especially in the BT549 cells, we made several probes that spanned this area and incubated them with nuclear extracts from the MCF-7 and BT549 cell lines. All of the fragments depicted in the schematic drawings of figures 9 and 10, including the fragments that are not represented by gel shifts, were capable of binding protein but the shifting pattern was the same with nuclear extracts from both cell lines. Figure 9 shows representative probes from region 2 and figure 10 shows representative probes from region 3.

Figure 11 depicts the gel shift assay for the 140bp XbaI-PvuII fragment from region 3 that showed enhancer activity with the SV40 promoter in BT549 cells. The pattern of shifted bands was the same for both MCF-7 and BT549 nuclear extracts. We also digested the 140bp fragment into smaller fragments and we did not see a difference in shifted patterns (figure 12). There was a consistent main band, and then depending on the particular assay there were an upper and a lower band or a series of bands. The upper and lower bands were usually more pronounced when the MCF-7 nuclear extract was used, but they could be seen with the BT549 nuclear extract with increased amount of extract or upon longer exposure to film.

Gel shift competitions were performed with all of the various combinations of the 140bp fragment and its subfragments. Only after adding 100X competitor could any reduction be seen in the main band produced by the shift whether the competition was performed with the 140bp fragment or its subfragments. The upper and lower bands are more sensitive to the competition since they were reduced by 50X and sometimes 10X competitor. The best competition we observed was with the 49bp XbaI-HinfI probe (figure 13). Due to these observations we had a 35mer synthesized that spans from the DdeI restriction site to the HinfI site. It encompasses all but 14bp of the 49bp XbaI-HinfI probe. The 35mer produced a shift of one band much like the main band seen in the previous shifts with both extracts, and effectively competes itself (figure 14). The 140bp XbaI-PvuII probe was competed with the 35mer and almost complete disappearance of the shifted band occurred with 50X competitor (figure 15). The 35mer was also capable of competing the main band in the shift with the 91bp HinfI-PvuII fragment. Even though the 91bp HinfI-PvuII fragment does not contain the 35mer fragment, it was almost completely competed with 50X competitor (figure 16). However, the 35mer and the 91bp HinfI-PvuII fragment do have spans of 4-6bp in common. The four base sequences are AAGA, TATG, and TTTC. The five base sequences are TGACT and GAGGT and the six base sequence is AGAGGT.

To observe whether or not the 140bp XbaI-PvuII fragment and the 35mer are capable of binding proteins in a similar manner in a cell line that was not derived from breast cancer, we performed the gel shift assay with HeLa cell nuclear extract (figure 17). HeLa cells have about twenty times the amount of EGFR than the MCF-7 cells (21). The 35mer probe produced a shift that has the same mobility as the shift seen with the breast cancer cell line nuclear extracts; however, the 140bp probe produced a higher shift when it was incubated with HeLa extract. It is possible that the same binding protein appears in both cell lines and binds to the 35mer sequence, but when the whole 140bp fragment is used as a probe a different complex of proteins that shifts higher is available in the HeLa cells.

Competitions were performed with other synthetic oligonucleotides that correspond to known cis elements to determine specificity and to see if this DNA-protein interaction involved any of these known transcription factors. The probe for the gel shift assay seen in figure 18 was the 140bp XbaI-PvuII fragment. There was very little competition with the SP1 oligonucleotide which was not surprising since it has no base combinations in common with the 35mer. The AP-2 oligonucleotide was capable of competing the main band at 50X competitor but not as well as 50X of the 35mer. Interestingly, there is no homology between the 35mer and the concensus AP-2 binding site; however, two four base sequences, ACTG and TGAC, match between the oligonucleotides (see figure 18). The NF-1 oligonucleotide has two four base sequences, ATTT and TATG, in the oligonucleotide and a five base stretch, GCCAA, that is within the concensus element in common with the 35mer, but it was unable to compete the shifted band with 50X competitor.

The previous gel shift competition assays led us to the realization that the sequence TGAC is the only sequence that the 35mer, the AP-2 oligonucleotide and the 91 HinfI-PvuII fragment have in common. Since the AP-2 oligonucleotide was capable of slightly competing the main shifted band (figure 18) and the 35mer was capable of competing away the shifted band observed in the gel shift assay with the 91 HinfI-PvuII probe (figure 16), we decided to scramble the bases ATGACT within the 35mer to see if these sequences were directly involved in the DNA-protein interaction. The mutant 35mer, with the scrambled bases, was unable to compete with the 140bp XbaI-PvuII probe for the protein involved in the main shifted band. The mutant 35mer was unable to compete when it was present in 100X excess unlike the complete competition observed with 50X of the wild type 35mer (figure 19B). When the mutant 35mer was used as a probe it was able to produce a shifted band with both nuclear extracts (figure 19C). The optical densities of the shifted bands produced with the wild type and the mutant 35mer were measured by densitometry and the shifted band of the mutant 35mer was only one eighth of the intensity of the shifted band seen with the wild type 35mer using the same amount of nuclear extract. Additionally, the mutant 35mer probe did not have the same mobility in the gel as the wild type probe. This suggests that the shift that is observed with the mutant probe could represent another binding site within the 35mer. However, the inability of the mutant 35mer to compete and to strongly bind protein is convincing evidence that the sequence ATGACT is involved in the binding site for the DNA-protein interaction represented by the main shifted band in these gel shift assays.

DNase I Protection Assay

The DNase I protection assay, or footprinting technique (38,39) delineates, at the nucleotide level, specific regions of DNA-protein interaction, and provides the sequence of the specific cis element that is involved in factor binding observed with the gel shift assays. End labeled probes corresponding to the enhancer region 3b were incubated with nuclear extracts from the MCF-7 and BT549 cells and then treated with DNase I (see Materials and Methods). This enzyme makes single stranded nicks on DNA and regions that are protected from protein do not become nicked. Therefore, a comparison of the DNase I digested

patterns between the probe alone and the DNA that has bound protein displays a region that is lighter, or protected from the action of the DNase I and hypersensitive bands that show areas where DNase I preferentially cleaves, usually adjacent to protected regions.

Figure 20 depicts the DNase I protection of the 35bp DdeI-HinfI fragment that is specific for the DNA-protein interaction seen with this region. We performed the DNase I protection of the 140bp XbaI-PvuII fragment and found that when the DNase I digested free probe was compared to the probe that had bound protein there were lighter regions, especially as the concentration of nuclear extract was increased, but it was difficult to determine specific regions that were significantly protected from the DNase I digestion. Therefore, we decided to use the 35bp DdeI-HinfI fragment as a probe anticipating that a smaller probe might help delineate a region of protection. Interestingly, two regions of protection appeared, but varied protection of the probe occurred among the different extract concentrations. Incubation of the probe with all of the concentrations of nuclear extract produced a region of protection surrounded by hypersensitive bands, region A, in the more 5' region of the probe (figure 20B). When the concentration of the nuclear extract was increased there was protection of another region, region B, in the more 3' region of the probe. This suggests that one protein is capable of specifically binding to the probe upon the addition of extract, and the increase in extract caused protein to complex onto another region of the DNA possibly due to the binding of the first protein. Based on a sequencing reaction that was electrophoresed alongside a DNase I protection assay with the 35mer probe, the protected region A corresponds to the sequence TATGA, and region B corresponds to the sequence CACCATTTC.

These results suggest that binding at region B is secondary to that at region A since mutation of the A region produced a fragment with very limited binding in a gel shift assay and that could not compete with the intact 140bp fragment. Additionally, a fragment corresponding to the 3' half of the 35mer (which contains region B but not region A) could not compete with the 140bp fragment for binding. We have preliminary results from transient transfection assays indicating that the wild type 35mer retains enhancer activity in BT549 cells (figure 21). The fold induction was 2.5X in conjunction with the SV40 promoter, and 9.3X with the EGFR promoter, indicating this element functions preferentially with its homologous promoter. With both promoters, the activity of the mutant 35mer was approximately half that of the wild type 35mer. These results suggest that region B (or other sequences in the 35mer) also have enhancer activity. It is possible that the limited binding we see by gel shift with the mutant 35mer in gel shift assays is stabilized through interactions with the EGFR promoter and thus has limited functionality.

To further examine the protein factors interacting with the intron repressor region, we are performing Southwestern analysis. Preliminary experiments with a panel of nuclear extracts from ER+ and ER- breast cancer cell lines suggest that this region can bind a factor approximately 130 kDa in size that is present all of the cell lines, as well as two factors approximately 45 and 32 kDa in size that are only seen in extracts from ER+ cells. Taken together with the results of the competition gel shifts described above, this suggests that the 130 kDa factor may bind to the enhancer element, and the 45 and/or 32 kDa factors may be involved in repression.

Micrococcal Nuclease Assay

Since gel shift data suggests that potential factors are in both cell types, one possible mechanism for the differential regulation of EGFR may involve the control of regulatory factor access to their binding sites by positioning of nucleosomes. To investigate if the position of nucleosomes differed in the EGFR first intron in low versus high EGFR expressing breast cancer cells, nuclei were isolated from MCF-7 and

BT549 cells. The nuclei were subjected to increasing concentrations of micrococcal nuclease and the DNA was purified for Southern hybridization and probed with a fragment corresponding to the 3' end of the intron region of interest. The micrococcal nuclease digests the linker region between nucleosomes allowing us to view the phasing pattern within the intron.

The resulting Southern blot can be seen in figure 22A. The solid arrows denote lanes for comparison between the MCF-7 and BT549 cell lines. The shaded arrow depicts a region of nucleosome disruption within the phasing of the EGFR gene in BT549 cells that cannot be seen in the MCF-7 cells. The open arrows indicate the region of an 80bp shift of nucleosomes in the BT549 cells perhaps due to disruption more 5' in the intron. Figure 22B shows a schematic model of the nucleosome phasing in the EGFR first intron in these low and high EGFR expressing cell lines. Also shown are the aforementioned DNase I hypersensitive sites. This schematic model was determined by measuring the distances between the molecular weight markers that were electrophoresed alongside the micrococcal nuclease assay and extrapolating the molecular weight of the nucleosomal bands. These nucleosomal bands were then counted and their distances were compared with the marker and with each other. The positioning of the nucleosomes relative to the DNase I hypersensitive sites in these two cell lines suggests that the nucleosomes may prevent binding of regulatory factors to the first intron in the MCF-7 cells, and the binding of a specific factor in the BT549 cells may alter this spacing and allow the binding of other common regulatory factors that result in the up-regulation of EGFR.

DISCUSSION

Overview

In order to investigate the regulatory role of the first intron of EGFR in breast cancer cells we examined the ability of the intron regions to act as effectors of CAT expression with both a heterologous and the homologous promoter. We further elucidated the capacity of the regions of the first intron to bind protein specifically through electrophoretic mobility shift assays and DNase I protection assays. Lastly, we examined the nucleosome phasing of the first intron in both low and high EGFR expressing breast cancer cell lines with the micrococcal nuclease assay. These techniques allowed us to elucidate an enhancer region that is specific to the high EGFR expressing cell line and a candidate sequence for protein binding in this region. We also discovered altered nucleosome phasing in the first intron of the high EGFR expressing cells.

Transcriptional Analysis

Our series of CAT constructs were utilized to investigate the ability of the first intron regions of EGFR to affect transcriptional activity and, therefore, expression of the CAT gene. The use of both a heterologous and the homologous promoter allowed us to examine which intronic elements are capable of affecting expression with a general promoter and which elements are able to affect expression in the presence of the natural EGFR promoter. Others have suggested that intron enhancer elements are stronger activators of transcription in concert with their homologous promoters than heterologous promoters (53,55). In agreement with these findings, our results with the 35mer enhancer element indicate that while this intronic regulatory element can function with a heterologous promoter, it has a strong prefernce for the homologous EGFR promoter.

In general, when the pSV-CAT series of constructs were transfected into the ER positive, low EGFR expressing cell lines they were not able to induce CAT activity to the same extent as in the ER negative, high EGFR expressing cell lines. In HeLa cells, a similar construct to our pSVI-CAT construct has been reported to enhance CAT activity 3 fold (21). By comparison, the intron region has the ability

to induce trancription almost twice as much in the EGFR overexpressing breast cancer cell lines. However, the difference in the fold induction of the CAT activity seen between breast cancer cell lines that are low versus high expressors was not significant until region 3 was investigated. The enhancer like activity of region 3 was more pronounced in the BT549 cells and was significantly different from the activity in the MCF-7 cells as observed in CAT assays with the constructs containing region 3 and region 3b. In comparing region 3a with region 3b it is clear that the ability of region 3 to act as an enhancer can be pinpointed to region 3b, a 140bp fragment of DNA. It is also clear from the CAT assays performed with pSVIΔ3b-CAT that region 3b is a stronger enhancer and essential to the regulatory ability of the intron region in the high EGFR expressing cell line. It has been observed by others that a transcription termination region that partially blocks RNA elongation in A431 cells is located just upstream of region 3b (45). Additionally region 3b is close to but not part of the intron enhancer that has activity in HeLa cells and shows several protein binding sites (21). We consider this 140bp region to be a crucial part of the regulatory ability of the intron region in breast cancer cella and to be an enhancer that is specific for the cells that overexpress EGFR.

The pJFE-CAT series of constructs was designed to examine the ability of regulatory intron regions to interact with their natural promoter. The use of the homologous promoter also insures that the regulatory ability of the intron regions relates to the in vivo situation. Others have found that in transient transfections the EGFR promoter has high activity relative to the SV40 early promoter or the Rous sarcoma virus LTR (11,16). Our pJFE-CAT had higher basal activity than the pSV-CAT which may contribute to the overall lower fold induction observed with the pJFE-CAT series of constructs. Between the cell lines there was no significant difference in the regulatory ability of the different regions of the intron in conjunction with the EGFR promoter. However, region 6 was capable of inducing activity in all of the cell lines, regardless of EGFR expression. It is possible that this region of DNA or a smaller portion of it is capable of regulating expression in a cell-specific manner, but since it is removed from its natural placement in the gene it may not be able to regulate specifically. Our recent experiments with the 35mer indicate that this sequence functions as an enhancer preferentially with the EGFR promoter. We are currently assessing the cell-specificity of this activity.

Maekawa et al. (21) found that the region that corresponds to our region 6 is capable of enhancing activity in HeLa cells when CAT gene expression is being driven by the EGFR promoter. They also found that when the region that corresponds to our region 3 was subcloned into the enhancer polylinker site of a construct that contained the EGFR promoter driving CAT expression that it had no activity in HeLa cells. From these experiments, they concluded that a region that is approximately 1.8kb downstream of the exon 1/intron 1 boundary functions as an enhancer in HeLa cells. This 136bp region can be found between the SstI and PstI restriction sites in region 2 (see figure 23). Due to our combined data with the $pJFEI_6$-CAT and the $pSVI_{3a}$-CAT constructs, we can also delineate this region as a potential enhancer-like region in breast cancer cell lines.

In a similar manner we have mapped the regulatory regions within the first intron of EGFR as observed in breast cancer cells. Figure 23 depicts our current understanding of the these regions as predicted from our accumulated CAT data. It is important to remember that endogenously these elements are adjacent and could have an effect on each other that will not be detected in a CAT assay. This regulatory map can act as a guide for our future studies with regulatory elements of the first intron of EGFR and for future mutagenesis studies of constructs in order to delineate which regulatory elements are essential for the differential expression of EGFR in breast cancer cell lines.

In Vitro Analysis

According to our DNase I hypersensitivity and CAT assays, the 2.1kb first intron region of EGFR has several elements that have the ability to bind factors and influence expression. We utilized three in vitro techniques to elucidate the ability of regulatory factors to bind to specific cis elements and to determine the sequence of the elements: electrophoretic mobility shift assays (or gel shift assays), DNase I footprinting and sequencing. The sequencing of the 2.8kb region of the first intron was instrumental in choosing restriction sites for the other assays and for the final delineation of the cis element.

Regions 2 and 3 of the EGFR first intron had the ability to induce CAT expression at least two fold in all of the cell lines tested, but there was no significant difference in the binding of protein between the low and high EGFR expressors. Longer exposure of the gels to film and/or higher concentrations of protein also showed shifted bands at the same regions of the gel. There is a possibility that further reduction in the size of the probes used for the gel shifts might elucidate a difference in binding patterns between the cell lines. For example, in region 2, the 200bp and 209bp probes are good candidates for closer examination.

Once the enhancer activity of region 3b was determined through CAT assays, we wanted to examine the ability of this 140bp XbaI-PvuII fragment to bind protein. When the 140bp region and its subdivided regions are used as probes for gel shifts the main shifted band, and even the bands that appear above and below the main band, shift to the same region on the gel with extracts from cells with different levels of endogenous EGFR expression. The secondary upper and lower bands appear with increasing concentration of protein and are presumed to represent complexes of proteins that are binding to the DNA-protein interaction that is represented by the main shifted band. The intensity of the appearance of these bands may be contingent upon different preparations of crude nuclear extract from the cells.

Even though we saw no distinct difference in the shifted patterns observed with the breast cancer cell extracts, we chose to investigate the characteristics of the main shifted band. When the 140bp region and its subfragments were used as probes and competitors it was observed that the upper and lower bands that appear were easier to compete than the main band, again suggesting that these bands correspond to complexes that are formed secondary to the main DNA-protein interaction. These competitions led to the synthesis of the 35mer that displays one band shift, and effectively competes with probes from the 140bp region and its subfragments for the protein involved in the DNA-protein interaction represented by the main shifted band. There is evidence that this 35mer is not as effective at competing the upper band observed in the gel shifts utilizing the 140bp probe, especially in the MCF-7 cells. This upper band could potentially represent another DNA-protein interaction that occurs somewhere else in the 140bp region and possibly only in the MCF-7 cells. The 35bp region appears to be the binding site for the protein involved in the DNA-protein interaction seen in the main shifted band, and our transient transfection experiments with the 35mer subcloned into the pSV-CAT and pJFE-CAT vectors indicate that this fragment retains the ability to enhance transcription when isolated from other intron sequences.

The gel shifts with the 140bp region and the 35mer using the HeLa cell nuclear extracts were performed to see if shifts occur that are similar to the ones observed in the breast cancer cells. The HeLa extracts contain a protein that binds to the 35mer probe and produces an identical shift; however, the 140bp region has reduced mobility in the gel when it is bound by proteins from the HeLa extracts as compared to the main band detected with the breast cancer cell line extracts. This suggests that the protein that binds to the 35mer may be ubiquitous, while the complexing of proteins on the 140bp region is different in the Hela cells as compared to the breast cancer cells.

The gel shift competitions performed with the common cis element oligonucleotides showed that the competition observed with the 35mer is specific. Even though the AP-2 oligonucleotide can compete with the 140bp probe for the protein, it is not a complete competition at 50X like the one seen with the 35mer, indicating that the 35mer is more specific. Interestingly, the combined use of the oligonucleotides and the 35mer in this competition assay provided valuable information on the protein binding site. The 35mer does not contain the consensus sequences of the SP1, AP-2 or NF-1 cis elements. However, the 35mer does share some stretches of bases in common with the flanking regions in the AP-2 and NF-1 oligonucleotides. Since the NF-1 oligonucleotide does not compete with the main shifted band those sequences can be disregarded. The AP-2 has two stretches of four base sequences in common with the 35mer and they are situated close to each other. The gel shift assay with the 91bp HinfI-PvuII fragment as the probe and the 35mer as competitor also provided important information. The 91bp HinfI-PvuII fragment is a portion of the 140bp region that does not contain the 35bp region, but the 35mer is capable of competing with this probe for the protein involved in its shifted band. The 91bp HinfI-PvuII fragment and the 35mer have sequences in common, but the only stretch of sequences that they have in common with the AP-2 oligonucleotide, which is capable of some competition, is TGAC.

We have also analyzed the sequence of the 35bp region for putative cis elements and found no direct matches with any of the cis elements in the GCG database. The closest match is the cis element for GCN4 which is TGAC/GTCAT. The yeast protein GCN4 induces transcription of several yeast genes involved in amino acid biosynthesis (56). Interestingly, our candidate sequence for the DNA-protein interaction we have observed is half of the GCN4 site. The DNA binding region of GCN4 shows strong homology at the amino acid level to v-jun, the oncogene of avian sarcoma virus ASV17 (57). It was through this homology that Jun was found to be one of the proteins that binds to the AP-1 cis element whose sequence is similar to the GCN4 cis element. The AP-1 cis element sequence, TGAGTCAG, is also somewhat different from the candidate binding site in the 35mer, TGAC. AP-1 does not appear as a putative binding site in the 35mer with or without mismatches using the GCG program.

We made a mutant 35mer that has the sequence ATGACT scrambled to CATGTA to examine the importance of these bases in the DNA-protein interactions observed with the gel shift assays. Previously, we have mentioned that the bases TGAC are four bases that seem important when we analyze the sequences of the probes and the competitors involved in our gel shifts. Since the competition with the AP-2 oligonucleotide was not complete, we chose to add a base on each side of this sequence in an attempt to elucidate a specific binding site. The lack of competition observed with the mutant 35mer in comparison with the complete competition observed with 50X of the wild type 35mer strongly indicates that we have elucidated the binding site for the DNA-protein interaction that produces the main shifted band in our gel shift assay. It appears as though another binding site exists in the 35mer sequence which accounts for the faint shift observed with the mutant 35mer. This is potentially a binding site for a different protein since the mutant probe which has a different mobility than the wild type probe produces a shift that has a similar mobility to that observed with the wild type probe. However, this other binding site in the 35mer seems secondary not only because of the lack of intensity in the observed shift with both nuclear extracts, but also because of the faint increase in binding with increasing nuclear extract. Further examination of the secondary binding sites within the 140bp XbaI-PvuII fragment and the 35mer through mutation analysis could elucidate a difference in the two breast cancer cell lines with regard to protein binding within these fragments.

Our DNase I protection assays show that protein from breast cancer cell nuclear extracts can protect two regions within the 35bp DdeI-HinfI fragment. When the entire 140bp XbaI-PvuII fragment is used

as a probe for this assay slight regions of protection can be visualized. However, when the smaller 35bp region that contains the major region of protein binding within the 140bp fragment is utilized as a probe distinct areas of protection are visable. Based on sequencing reactions that were electrophoresed alongside the protection assay, we concluded that region A corresponds to the sequence TATGA which overlaps with the candidate binding site ATGACT that was determined through the gel shift assays. Another protected region that occurs with increased nuclear extract concentrations is region B which corresponds to the sequence CACCATTTC. It is possible that the DNA-protein binding that occurs at this element in the presence of more extract might be capable of binding protein only after region A has protein bound. Other evidence that leads to this conclusion is the inability of the sequence ATTT from the NF-1 oligonucleotide to compete with the binding represented by the main shifted band in figure 18. However, these bases within region B may not be essential for protein binding. Our transient transfection assays with the wild type and mutant 35mers suggest that region B also has some enhancer activity. It is possible that the limited binding we see by gel shift with the mutant 35mer in gel shift assays is stabilized through interactions with the EGFR promoter and thus has limited functionality. Alternatively, the 6 bases that were scrambled in the mutant 35mer may leave enough of region A intact to allow for some protein interaction that again would be stabilized in the presence of the EGFR promoter.

In Vivo Analysis

The previously mentioned assays investigated the regulatory regions of the first intron of EGFR in an in vitro environment. In other words, isolated regions of the intron are placed in an artificial situation to examine their ability to regulate transcriptional activity or their ability to bind proteins sequence specifically. These are useful methods that provide us with valuable information. However, they do not enable us to observe the interactions of the larger regulatory unit consisting of the promoter, exon 1, and intron 1 of the EGFR gene, in which promoter-intron interactions may be critical and the spacing of elements in these regions relative to each other may be important in the differential regulation of EGFR. Therefore, we designed a "native" EGFR reporter construct that maintains intact the relative spacing of the EGFR elements. This construct consists of a 3.7kb fragment containing an uninterrupted stretch of the EGFR promoter (~1kb), exon 1, and intron 1 (~2.5kb), upstream of a β-globin reporter gene. This construct was transiently transfected into MCF-7 and BT549 breast cancer cells, and the level of β-globin transcript from the reporter gene was measured by RNase protection and normalized to the endogenous level of 36B4 mRNA. Initial experiments showed a strong β-globin signal for the BT549 cells, and a greatly reduced signal for the MCF-7 cells. These findings generated considerable excitement since they lent strong support to our hypothesis that the intact structure of the EGFR promoter/exon 1/intron 1 region is important for differential regulation in ER+ vs. ER- breast cancer cells. However, while these results were highly consistent in subsequent assays, we observed that the control, promoterless β-globin reporter construct unpredictably gave a signal in some experiments. Despite consultations with the creator of this β-globin reporter plasmid and several attempts to determine the cause of this unexpected expression, it remains a mystery. With our negative control giving us unexplained, unpredictable positive results, we could not trust the results with our native construct. Therefore, we obtained different β-globin reporter plasmids and tested them extensively for lack of expression in the promoterless state.

The pGLOB3 reporter construct from Dr. Meinrad Busslinger (59) consistently showed no expression in MCF-7 and BT549 cells. We have now subcloned the "native" EGFR insert into this reporter construct, and have begun transient transfection assays in MCF-7 and BT549 cells. Our preliminary results show differential expression of this reporter that reflects the level of the endogenous EGFR gene in these cells. Once these results are confirmed, we will mutate or delete specific sites

(starting with the enhancer element) and combinations of sites in the EGFR insert and compare the effect on the level of the β-globin reporter gene to what we have observed with our conventional CAT reporter constructs. For example, with the enhancer element, we will first substitute the mutated 35mer for the wild type sequence. Future experiments will be based on our ongoing characterization of the various EGFR elements and will likely include deletion of the 3' half of the 35mer sequence (leaving only region A), mutation of both region A and B, and combinations of these with specific mutations in promoter and repressor elements as they are delineated. Additionally, we will perform stable transfections in MCF-7 and BT549 cells with the parental native construct and select variants (co-transfected with a plasmid containing the neomycin resistance gene), to assess the role of chromatin structure in the differential regulation of EGFR.

Another way of investigating the natural configuration of the EGFR gene is with in vivo approaches such as the DNase I hypersensitivity assay and the micrococcal nuclease assay. These methods allow us to visualize the accessibility of the regulatory regions of the gene to DNA nicking or cutting enzymes which, in turn, reflects the openess of the chromatin structure of a region of the gene. Evidence has shown that promoters and enhancers are typically marked by hypersensitive domains in chromatin that reflect the absence of canonical nucleosomes (58). Therefore, we utilized the micrococcal nuclease assay to investigate the nucleosome phasing of the first intron of the EGFR. We found that the first intron of EGFR had regular phasing of nucleosomes in the low EGFR expressing MCF-7 cells, and that the first intron of EGFR had altered phasing of nucleosomes in the high EGFR expressing BT549 cells by comparison. The micrococcal nuclease assay in BT549 cells exhibited a smearing of nucleosomal patterns, a disrupted nucleosome and an 80bp shift that could potentially be an effect of the disrupted nucleosome. It is possible that this disrupted nucleosome opens up the EGFR intron 1 for enhancer proteins to reach cis elements in the BT549 cells while the regularly spaced nucleosomes observed on the first intron of EGFR in MCF-7 cells do not allow regulatory factors to access these same elements. This suggests that the first intron of EGFR in the BT549 cells, or an EGFR overexpressing cell line, has an open conformation that could allow regulatory factors to access DNA and influence expression of the gene. The results of this assay are in agreement with our previous observations using the DNase I hypersensitivity assay (22).

A question that arises is whether or not our transiently transfected $pSVI_{3b}$-CAT construct has nucleosome phasing that may account for its enhancer activity in the high EGFR expressing cell line. Others have shown that nucleosome phasing can occur on transiently transfected DNA. SV40 minichromosomes and bovine papilloma virus based episomes are not integrated into the genome and are capable of nucleosome phasing that is similar to that of integrated copies (60-62). Other pBR322 and pUC based transiently transfected plasmids have shown weak nucleosome phasing that is not representative of integrated plasmid copies (61). Therefore, our pUC based transiently transfected plasmids could be able to exhibit nucleosome phasing, but it might not be indicative of the in vivo situation. We could test our constructs for nucleosome phasing, but we propose that the enhancing activity of region 3b in our CAT assay system is more likely due to the interaction of this region with a co-activator.

Further use of in vivo techniques could better elucidate the role of the first intron in the differential expression of EGFR in breast cancer cells. All of these techniques involve isolation of nuclei from the cell, treatment of the nuclei, purification of the DNA, and electrophoresing the DNA to determine the size of the DNA product. Another way to investigate whether or not a nucleosome is displaced in the first intron in the BT549 cells is to perform an endonuclease accessibility assay. In this assay, nuclei are exposed to endonucleases for restriction sites that are within the presumed region of nucleosome binding. If the site is protected by a nucleosome then the resulting DNA fragment will be larger than the DNA product that

will occur if the restriction enzyme has access to this site. Alternatively, using a combination of the amplified primer extension (APEX) (63) and native genomic blotting techniques (23,24) we can determine the precise sites of protein binding within the first intron regions of EGFR with in vivo footprinting. In this assay, DNase I cleaved DNA is produced by treatment of nuclei with DNase I as in the hypersensitivity assay. The DNA is then digested with a restriction enzyme that yields a fragment of 1-2kb. The DNA is hybridized to a synthetic oligonucleotide primer that is complimentary to one end of the fragment, and subjected to repeated rounds of extension. This APEX procedure results in a linear amplification of fragments in which one end corresponds to the primer and the other end to sites of DNase I digestion. The DNA is then blotted and probed, and the size of the amplified fragments indicates the regions within the intron that are binding protein in vivo.

A third approach is the nuclear ligation assay (64), which could be used to determine if the DNA looping model applies to the potential interaction between the promoter and first intron regions of EGFR. In this case nuclei are treated with a restriction enzyme that would produce a DNA fragment that contains the promoter and the first intron region of interest. Then the nuclei are treated with DNA ligase and purified DNA is cleaved at a restriction site within the middle of the DNA fragment and PCR is performed with primers from either end of the original DNA fragment. Analysis of the PCR products determines if ligation occured, and successful ligation indicates that the ends of the DNA fragment are in close proximity in vivo. This proximity of the ends of the DNA could be mediated by tethering proteins or direct protein-protein interactions. This would imply that there is DNA looping to allow the basal transcription factors to interact with the trans-acting factors in the first intron.

CONCLUSIONS

We have investigated the regulatory role of the first intron in the expression of EGFR and have identified a cis element that could be involved in the overexpression of this gene in ER negative, high EGFR expressing breast cancer cells. The 2.5kb region of the first intron of EGFR seems to have regions of both positive and negative regulatory ability. In general, the regions of the intron had similar activity in both low and high EGFR expressing cell lines until an enhancer region, region 3, was elucidated. A smaller portion of this region, region 3b, was capable of enhancing transcriptional activity specifically in the high EGFR expressing cell line and it was found to be essential to the regulatory ability of the 2.1kb intron region in these cells. Further examination of the region 3b determined that a 35bp region contains the major site of DNA-protein interaction for this 140bp region and retains enhancer activity. The proteins that bind to the 35bp region can be found in low and high EGFR expressing cell lines, and protection assays indicate two regions of protein binding. A candidate sequence for the cis element involved in this interaction is ATGACT. Inspection of this sequence indicates that it may bind a novel protein that is in direct contact with the DNA. It is possible that co-factors are responsible for the use of this element as an enhancer in the high EGFR expressing cell line. Another protein may interact with the DNA-protein complex formed on this element to repress its activity in MCF-7 cells, or to help activate its regulatory role in BT549 cells to enhance the expression of EGFR. We propose that protein-protein interactions are involved in the utilization of this element to regulate the overexpression of the EGFR gene.

The in vivo analysis of the first intron of the EGFR gene led us to believe that this region, in low EGFR expressors, has regularly phased nucleosomes and, therefore, a closed conformation. However, in the high EGFR expressors, the first intron of EGFR exhibits disruption and displacement of nucleosomes. We hypothesize that the difference in EGFR expression between high and low EGFR expressing breast cancer cell lines may also be due to the accessibility of cis elements in the first intron to their cognate factors. The basic DNA elements have the ability to affect transcriptional activity when they are examined

in vitro in different cell types. The factors that act through these cis elements appear to be ubiquitously available. However, we propose that it is the configuration of the regulatory region of EGFR in vivo that determines the expression level of the gene.

Our studies have demonstrated that the 2.5kb portion of the first intron of EGFR has a complicated regulatory role and that there are several elements within the intron that may be involved in general expression as well as specific overexpression in hormone independent breast cancer cells. We have only just begun to understand and elucidate these regulatory elements and their context in vivo. Further work to comprehend the regulation of the expression of EGFR in breast cancer will hopefully lead to the ability to block the overproduction of EGFR with the intention of abolishing the rapid proliferation of hormone independent breast cancer.

REFERENCES

1. Fitzpatrick SL, Brightwell J, Wittliff J, Barrows GH, Schultz GS 1984 Epidermal growth factor binding by breast tumor biopsies and relationship to estrogen and progestin receptor levels. Cancer Res 44:3448-3453.
2. Sainsbury JRC, Farndon JR, Sherbert GV, Harris AL 1985 Epidermal growth factor receptors and oestrogen receptors in human breast cancer. Lancet i:364-366.
3. DiMarco E, Pierce JH, Fleming TP, Krauss MH, Molley CJ, Aaronson SA, DiFiora PP 1989 Autocrine interaction between TGF-a and EGF-receptor: quantitative requirement for induction of the malignant phenotype. Oncogene 4: 831-838.
4. Derynck R, Goddel DV, Ullrich A, Gutterman JU, Williams RD, Bringman TS, Berger WH 1987 Synthesis of mRNAs for TGF-α and β and epidermal growth factor receptor by human tumors. Cancer Res 47:707-712.
5. Gullick W J 1991 Prevalence of aberrant expression of the epidermal growth factor in human cancers. Br Med Bull 47:87-98.
6. Davidson NE, Gelmann EP, Lippman ME, Dickson RB 1987 Epidermal growth factor receptor gene expression in estrogen receptor-positive and negative human breast cancer cell lines. Mol. Endocrinol 1:216-223.
7. Bolufer P, Miralles F, Rodriguez A, Vasquez C, Lluch A, Garcia-Conde J, Olmos T 1990 Epidermal growth factor receptor in human breast cancer: coorelation with cytosolic and nuclear estrogen receptors and with biological and histological tumor characteristics. Eur J Cancer 26:283-290.
8. Nicholson S, Sainsbury JRC, Halcrow P, Chambers P, Farndon JR, Harris AL 1989 Expression of epidermal growth factor receptors associated with lack of response to endocrine tharapy in recurrent breast cancer. Lancet i:182-184.
9. Nicholson S, Richard J, Sainsbury C, Halcrow P, Kelly P, Angus B, Wright C, Henry J, Farndon JR, Harris AL 1991 Epidermal growth factor receptor; results of a 6 year follow-up study in operable breast cancer with emphasis on the node negative subgroup. Br J Cancer 63:146-150.
10. Koenders PG, Beex LVAM, Geurts-Moespot A, Heuvel JJTM, Kienhuis CBM, Benraad TJ 1991 Epidermal growth factor receptor-negative tumors are predominantly confined to the subgroup of estradiol receptor-positive human primary breast cancers. Cancer Res 51:4544-4548.
11. Haley J, Whittle N, Bennett P, Kinchington D, Ullrich A, Waterfield M 1987 The human EGF receptor gene: Structure of the 110kb locus and identification of sequences regulating its transcription. Oncogene Res 1:375-396.
12. Ishii S, Xu YH, Stratton RH, Roe BA, Merlino GT, Pastan I 1985 Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. Proc Natl Acad Sci USA 82:4920-4924.
13. King CR, Kraus MH, Williams LT, Merlino GT, Pastan IH, Aaronson SA 1985 Human tumor cell lines with EGF receptor gene amplification in the absence of aberrant sized mRNA. Nucl Acids Res 13:8477-8486.
14. Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Liberman TA, Schlessinger J, Downward J, Mayes ELV, Whittle N, Waterfield MD, Seeburg PH 1984 Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature 309:418-425.
15. Merlino GT, Johnson AC, Kageyama R, Pastan I 1989 Isolation and characterization of DNA binding factors regulating transcription of the EGF receptor proto-oncogene. In: Growth regulation of cancer II: UCLA symposia on molecular and cellular biology. (Lippman ME, Dickson RB, eds.), Alan R Liss, Inc, New York.

16. Johnson A, Ishii S, Jinno Y, Pastan I, Merlino GT 1988 Epidermal growth factor receptor gene promoter: Deletion analysis and identification of nuclear protein binding sites. J Biol Chem 263:5693-5699.
17. Kageyama R, Merlino GT, Pastan I 1988 Epidermal growth factor (EGF) receptor gene transcription: Requirement for Sp1 and an EGF receptor specific factor. J Biol Chem 263:6329-6336.
18. Kageyama R, Pastan I 1989 Molecular Cloning and Characterization of a Human DNA Binding Factor that Represses Transcription. Cell 59:815-825.
19. Maekawa T, Imamoto F, Merlino GT, Pastan I, Ishii S 1989 Cooperative function of two separate enhancers of the human epidermal growth factor receptor. J Biol Chem 264:5488-5494.
20. Gross DS, Gullick WT 1988 Nuclease hypersensitive sites in chromatin. Ann Rev Biochem 57:159-197.
21. Maekawa T, Imamoto F, Merlino GT, Pastan I, Ishii S 1989 Cooperative function of two separate enhancers of the human epidermal growth factor receptor proto-oncogene. J Biol Chem 264:5488-5494.
22. Chrysogelos, SA 1993 Chromatin structure of the EGFR gene suggests a role for intron 1 sequences in its regulation in breast cancer cells. Nucl Acids Res 21:5736-41.
23. Pauli U, Chrysogelos S, Stein J, Stein G 1988 Native genomic blotting: high resolution mapping of DNase I hypersensitive sites and protein-DNA interactions. Proc Natl Acad Sci USA 85:16-20.
24. Pauli U, Chrysogelos S, Stein J, Stein G 1988 Native genomic blotting: a novel approach to mapping of DNase I hypersensitive sites and protein-DNA interactions at high resolution. Biotechniques 6:142-146.
25. Fridovich-Keil JL, Gudas JM, Bryan IB, Pardee AB 1991 Improved expression vectors for eukaryotic promoter/enhancer studies. Biotechniques 11:572-579.
26. El-Ashry D, Chrysogelos SA, Lippman ME, Kern FG 1996 Estrogen-induction of TGF-α is mediated by an estrogen responsive element composed of two imperfect palindromes. J Steroid Biochem Mol Biol 59:261-269.
27. Wigler M, Pellicer A, Silverstein S, Axel R 1978 Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell 3:725-731.
28. Graham FL, Eb AJ van der 1973 A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456-467.
29. Hawley-Nelson P, Ciccarone V, Gebeyehu G, Jessee, Felgner PL 1993 Lipofectamine reagent: a new, higher eficiency polycationic liposome transfection reagent. Focus 15:73-83.
30. Hirt B 1967 Selective extraction of polyoma DNA from infected mouse cell cultures. J Mol Biol 26:365-369.
31. Murphey JM, Chrysogelos SA 1995 Direct measurement of transfection efficiency in transient transfection assays. BioTechniques 18:968-969.
32. Neumann, JR, Morency, CA, Russian, KO 1987 A novel rapid assay for chloramphenicol acety;transferase gene expression. Biotechniques 5:444-447.
33. Gorman CM, Moffat LF, Howard BH 1982 Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol Cell Biol 2:1044-1051.
34. Prost E, Moore DD 1986 CAT vectors for analysis of eukaryotic promoters and enhancers. Gene 45:107-111.
35. Garner M, Revzin A 1981 A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of Escherichia Coli lactose operon regulatory system. Nucl Acids Res 9:3047-3060.

36. Fried M, Crothers DM 1981 Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucl Acids Res 9:6505-6525.
37. Dignam JD, Lebovitz RM, Roeder RG 1983 Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucl Acids Res 11:1475-1489.
38. Galas D, Schmitz A 1987 DNase footprinting: A simple method for the detection of protein-DNA binding specificity. Nucl Acids Res 5:3157-3170.
39. Riegel AT, Remenick J, Wolford RG, Berard DS, Hager GL 1990 A novel transcriptional activator (PO-B) binds between the TATA box and cap site of the pro-opiomelanocortin gene. Nucl Acids Res 18:4513-4521.
40. Moreno ML, Chrysogelos SA, Stein G, Stein J 1986 Reversible changes in the nucleosomal organization of a human H4 histone gene during the cell cycle. Biochemistry 25:5364-5370.
41. Dynan WS, Tjian R 1985 Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316:774-778.
42. Ransone LJ, Verma IM 1990 Nuclear proto-oncogenes fos and jun. Annu Rev Cell Biol 6:539-557.
43. Imagawa M, Chiu R, Karin M 1987 Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell 51:251-260.
44. Santoro C, Mermod N, Andrews PC, Tjian R 1988 A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature 334:218-224.
45. Haley JD, Waterfield MD 1991 Contributory effects of de novo transcription and premature transcript termination in the regulation of human epidermal growth factor receptor proto-oncogene RNA synthesis. J Biol Chem 266:1746-1753.
46. Amundadottir LT, Johnson MD, Merlino G, Smith GH, Dickson RB 1995 Synergistic interaction of transforming growth factor alpha and c-myc in mouse mammary and salivary gland tumorigenesis. Cell Growth Differ 6:737-748.
47. Yarden RI, Lauber AH, El-Ashry D, Chrysogelos SA 1996 Bimodal regulation of epidermal growth factor receptor by estrogen in breast cancer cells. Endocrinology 137:2739-2747.
48. Vickers PJ, Dickson RB, Shoemaker R, Cowan KH 1988 A multidrug-resistant MCF-7 human breast cancer cell line which exhibits cross-resistance to antiestrogens and hormone-independent tumor growth in vivo. Mol Endocrinol 2:886-892.
49. Filmus J, Pollak MN, Cailleau R, Buick RN 1985 MDA-468, a breast cancer cell line with a high number of epidermal growth factor (EGF) receptors, has an amplified EGF receptor gene and is growth inhibited by EGF. Biochem Biophys Res Commun 128:898-905.
50. Horwitz, KB, Mochus, MB, Lessey, BA 1982 Variant T47D human breast cancer cells with high progesterone receptor levels despite estrogen and anti-estrogen resistance. Cell 28: 633-642.
51. Lu, J 1992 Magnesium chelation inactivates β-gal in -20°C storage. BioTechniques 12:177-181.
52. Farr A, Roman A 1991 A pitfall of using a second plasmid to determine transfection efficiency. Nucl Acids Res 20:920.
53. Franklin GC, Donovan M, Adam GI, Holmgren L, Pfeifer-Ohlsson S, Ohlsson R 1991 Expression of the human PDGF-B gene is regulated by both positively and negatively acting cell type-specific regulatory elements located in the first intron. EMBO J 10:1365-1373.
54. Rossi P, de Crombrugghe B 1987 Identification of a cell-specific transcriptional enhancer in the first intron of the mouse alpha 2 (type I) collagen gene. Proc Natl Acad Sci USA 84:5590-5594.
55. Beenken SW, Karsenty G, Raycroft L, Lozano G 1991 An intron binding protein is required for transformation ability of p53. Nucleic Acids Res 19:4747-4752.
56. Latchman, DS 1995 Eukaryotic transcription factors. Academic Press, San Diego, CA.
57. Vogt PK, Bos TJ, Doolittle RF 1987 Homology between the DNA-binding domain of the GCN4

regulatory protein of yeast and the carboxyl-terminal region of a protein coded for by the oncogene jun. Proc Natl Acad Sci USA 84:3316-3319.

58. Felsenfeld G 1992 Chromatin as an essential part of the transcriptional mechanism. Nature 355:219-224.
59. Bergers G, Graninger P, Braselmann S, Wrighton C, Busslinger M 1995 Transcriptional activation of the fra-1 gene by AP-1 mediated by regulatory sequences in the first intron. Mol Cell Biol 15:3748-3758.
60. Blasquez V, Stein A, Ambrose C, Bina M 1986 Simian virus 40 protein VP1 is involved in spacing nucleosomes in minichromosomes. J Mol Biol 191:97-106.
61. Jeong SW, Stein A 1994 DNA sequence affects nucleosome ordering on replicating plasmids in transfected COS-1 cells and in vitro. J Biol Chem 269:2197-2205.
62. Ostrowski MC, Richard-Foy H, Wolford RG, Berard DS, Hager GL 1983 Glucocorticoid regulation of transcription at an amplified, episomal promoter. Mol Cell Biol 3:2045-2057.
63. Cartwright IL, Kelly SE 1991 Probing the nature of chromosomal DNA-protein contacts by in vivo footprinting. Biotechniques 11:188-203.
64. Cullen KE, Kladde MP, Seyfred MA 1993 Interaction between transcription regulatory regions of prolactin chromatin. Science 261:203-205.

1	GTAAGGGCGT	GTCTGCGGCT	CCCGCGCCGC	CCCCGGATCG	CGCCCCGGAC	CCCGCAGCCC
61	GCCCAACCGC	ACACCGCGCA	CCGCTCGCGC	CCGCGCCCCG	CCCGTCCTTT	CCTGTTTCCT
121	TGAGATCACG	TGCGCCGCCG	ACCGGGACCG	CGGGAGGAAC	GGGACGTTTC	GTTCTTCGGC
181	CGGGAGAGTC	TGGGGCGGGC	GGAGGAGGAG	ACGCGTGGGA	CACCGGGCTG	CAGGCCAGGC
241	GGGGAACGGC	CGCCGGACCT	CCGCCCCGAA	CCGTCCAACT	TTCTTCCCTA	CTTTCCCCGC
301	CCAGCTGCGC	AGGATCGGCG	TCAGTGGGCG	AAAGCCGGGT	GCTGGTGGGC	GCCTGGGGCC
361	GGGGTCCGCA	CGTGCGCCCC	GCGCTGTCTT	CCCAGGGCGC	GACGGGGTCC	TGGCGCGCAC
421	CCGAGGGCGC	CGCTGCCCAC	CCGCCGAGAC	TGCACTGTTT	AGGGAAGCTG	AGGAAGGAAC
481	CCAAAAATAC	AGCCTCCCCT	CGGACCCCGC	GGGACAGGCG	GCTTTCTGAG	AGGACCTCCC
541	CGCCTCCGCC	TCCGCGCAGG	TCTCAAACTG	AAGCCGGCGC	CCGCCAGCCT	GGCCCCGGCC
601	CCTCTCCAGG	TCCCCGCGAT	CCTCGTTCCC	CAGTGTGGAG	TCGCAGCCTC	GACCTGGGAG
661	CTGGGAGAAC	TCGCTCTACC	ACCACTGCGC	TCCCGGGGAG	GGGTGGTGCT	GGCGGCGGTT
721	AGTTTCCTCG	TTGGCAAAAG	GCAGGTGGGG	TCCGACCCGC	CCCTTGGGCG	CAGACCCGGC
781	CGCTCGCCTC	GCCCGGTGCG	CCCTCGTCTT	GCCTATCCAA	GAGTGCCCCC	CACTCCCGGG
841	ACCCCAGCTC	CCTCCTGGCC	CGCGCCGAAA	CCCCAGGCTC	TCCTTCGATG	GCGCTCGCGG
901	AGACGTCCGG	GTCTGCTCCA	CCTGCAGCCT	TCGGTCGCGC	CTGGGCTTCG	CGGTGGAGCG
961	GACGCGGTGT	CCGGCCACTG	CAGGGGGGGA	TCGCGGACTC	TTGAGCGGAA	GCCCCGGAAG
1021	CAGAGCTCCA	TCCTGGCCAA	CACCATGGTG	TTTCAAAATG	GGGCTCACAG	CAAACTTCTC
1081	CTCAAAACCC	GGAGACTTTC	TTTCTTGGAT	GTCTCTTTTT	GCTGTTTGAA	GAATTTGAGC
1141	CAACCAAAAT	ATTAAACCTG	TCTTACACAC	ACACACACAC	ACACACACAC	ACACACACCG
1201	GATTGCTGTC	CCTGGTTCAA	GTGTGCCAAG	TGTGCAGACA	GAACATGAGC	GAGTCTGGCT
1261	TCGTGACTAG	CGACCATAAA	CCCACTTGAC	AGGGGAAACA	TGCCTTGGAA	GGTTTAATTG
1321	CACAATTCCA	ACCTTGAGTC	TGCGCGGGTT	CCAAGAGCCA	GGCCCGTACT	TGCTGTTGAT
1381	GTCATTGGCT	TGGGGAGTTG	GGGTTTGGTG	CCCGAGCGCG	GTCGTTGGGG	GAGGGCAAGG
1441	CATAGAACAG	TGGTTCCCAG	ACCCTTGCTG	CACATTGGAA	TTACCTGGGA	TTAAAAAAAA
1501	AAAATCAAAA	CAAAAACCAG	TGTCTCGTCC	CGCCCCCAGA	CATTCTGATT	TAATTGGCAT
1561	GGGGCAAGAC	CTGGACTTGG	GATTTTTTTT	AATGCTCTTC	ATGTGATCTG	TTGGGCAGCC
1621	AGATTTGGGG	ATCACTAGAC	GGAAGAAGGA	TTGTTAAAGT	CTCCGGAGAT	GTTACTTGCC
1681	AATGCTAAGA	GCTCTTTGAG	GACATCTGGA	ATTGTTACAA	TATTGCCAAA	TATAGGAAAG
1741	AGGGGAAAAG	GTAGAGTGTG	ATTCCAATAA	TAAAGGATTC	CGCTTTTCAT	TGAAGGAACT
1801	GGTGGAAAGG	TTTCTTCTCT	GCTGAGCCTG	CAGGCCCGTC	CTGCCTGCCT	GGGGTGCCCG
1861	GGAGCGGCCT	GCTCCGGAGA	CTGCTGACTG	CCGGTCCTGT	TAGTCAGGTG	TGAGCCCTGT
1921	CTCTGCCGAA	GAGACTCTTC	TCTTTATTTT	AAATTAAACC	CTCAGAGCAC	CACCAAAGCA
1981	TCACTTTTCT	CCCTCCATTG	GTGTTCTCAT	TCTTTGATGT	TACTTGTTTG	AACACCACTA
2041	TTAGTAGTTG	GAGATTTGTT	CCTGAGAAAA	ATATAAATAC	CACTTAATTT	GCCTGTTTGT
2101	CCCGCATTCA	CTCAAAACAG	AATGCTCCTG	CTAAGAGAAG	AGAGAGAGTA	GGAGAACAGA
2161	CGCTATTCCA	TTACAGTAAC	ATAAAACTGG	ATTTTCAGGG	GGACAAATTA	TTAAAATAGG
2221	AGATGAGCTC	TTTTAACAGA	AATTTGTTTA	AGGCCTGTGT	CTATCAAATT	CAGTGGATTT
2281	TATTCAAGAT	GCACTTTGTT	TAGTGGGAGT	TTTGTTTGGT	TCTGGGACAT	GCTAACTTCT
2341	AGACTTGCTG	CTCTTAGAGG	TTATGACTGC	CAAGACACCA	TTTCATGAGT	CCTAATCCCC
2401	ACATTAAGCA	TAAGAGGTGC	ACACTCTCCT	CCTATGGGGG	AAACTGAGGT	ACGAAGAACT
2461	AAAGTGACTT	TCCCACAGCT	GGTGGGAGGC	AGACGGGAAA	TTCACACCAG	GGGCTTCCAA
2521	CTCCAGATCC	CTCTCTCAAC	TTCCAAACTC	CACTGCCTTG	TCCGAGTTCT	GGTTTCAGGA
2581	GATCCAAATC	AGGTGTGTGC	AAATGTCTAA	TGTCAGAGCT	GGCAAGGGGA	AAGGGCCCAG
2641	GGAGCCGGGC	TCATGACGAT	GAGCCTGTCT	GAAGCTTCAA	CGGCGGGCTG	TCCGGCAGTC
2701	TGATTCCTGC	CGAGTTCCTC	AGCCCTCTGT	TGGGTCACCT	TCCATAGAGG	CAGCTTAGTC
2761	CTCAGTTCAG	TGAGCATGGA	GTGGAGACTG	CTTGAGGGGT	GCTGAGCAAA	GCCCTGCCTC
2821	TTACAGGATG	AAGGTGCTCT	CCAGAAGGGA	CACAGGAAAG	TATCAAGGCG	AGTCGAATTC

Figure 1. Sequence the 2.8kb region of the first intron of the EGFR gene spanning from the exon 1/intron 1 boundary to a downstream EcoRI site.

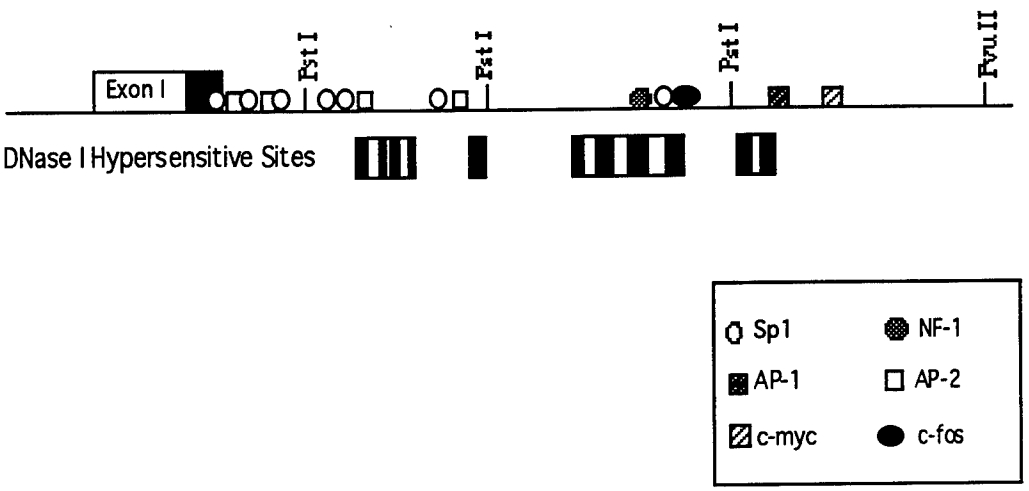


Figure 2. Location of putative transcription factor binding sites in the EGFR first intron. The c-myc site and the c-fos half site are consensus elements that were found in the sequence of the upstream 2.8kb region of the first intron of EGFR from the human liver cells and were not found in the corresponding sequence of the A431 cells.

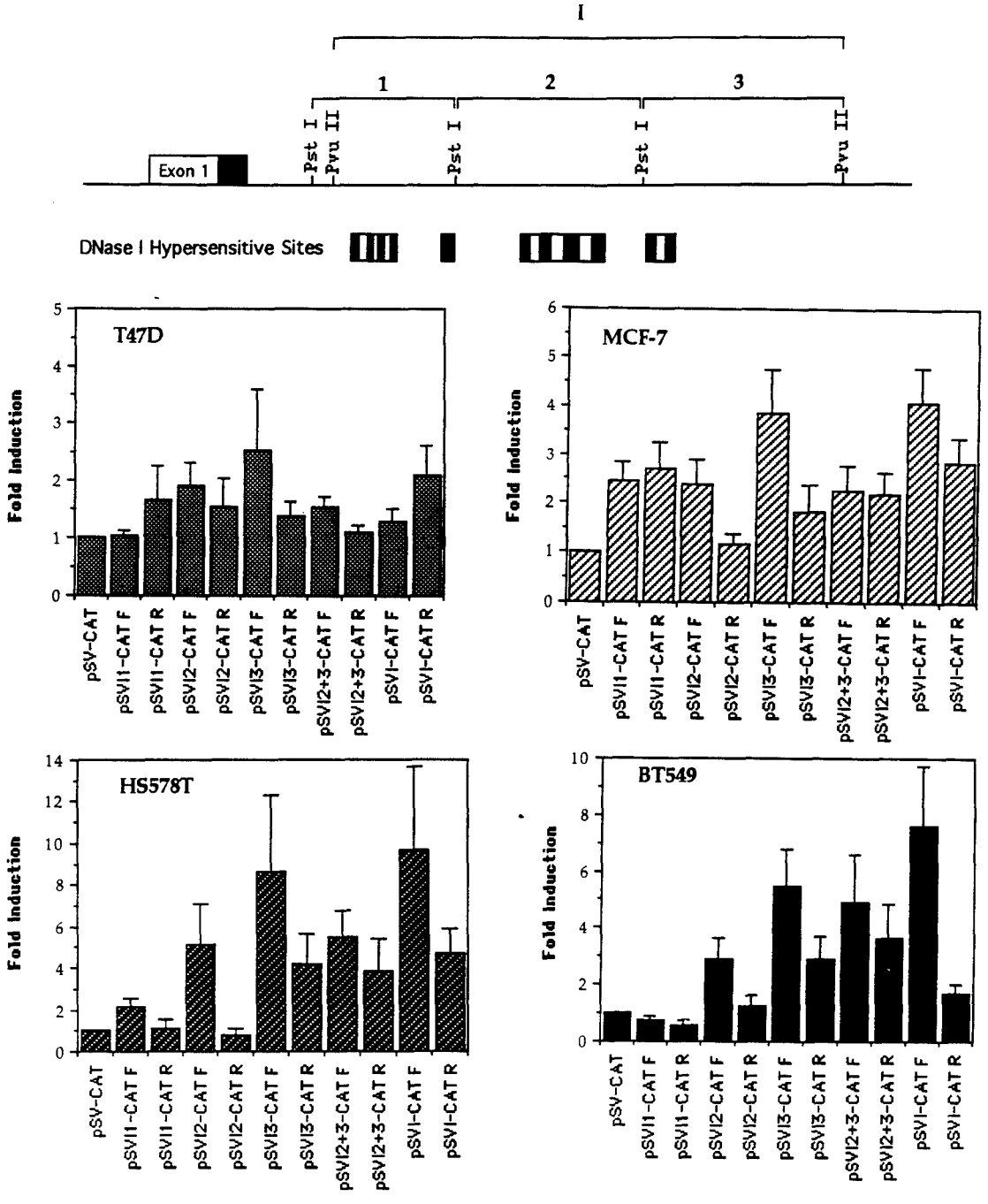
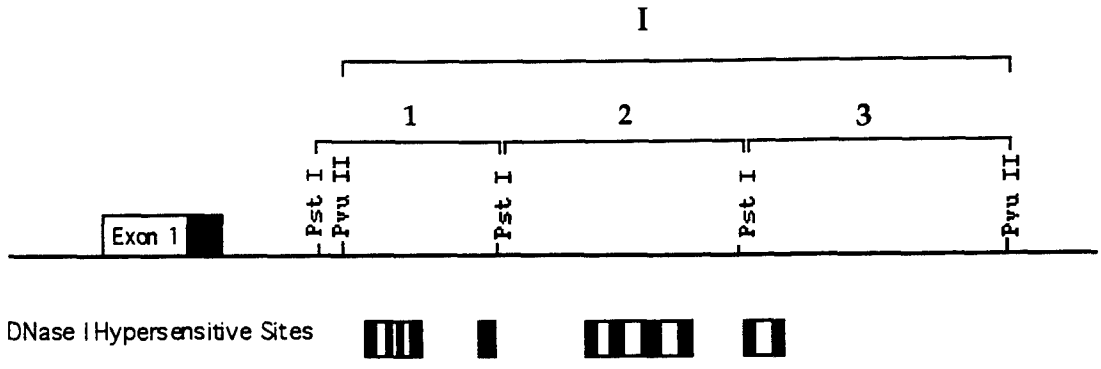


Figure 3. Relative CAT activity of the SV40 promoter-EGFR intron series of constructs in both the forward and reverse orientations in the MCF-7, T47D, HS578T and BT549 breast cancer cell lines.



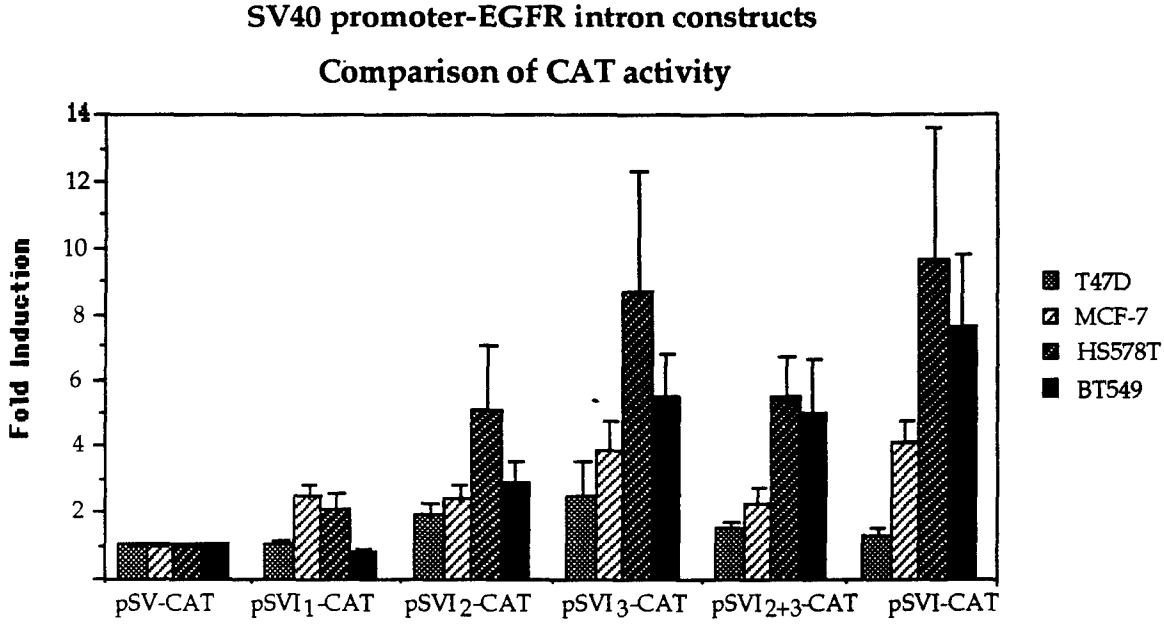
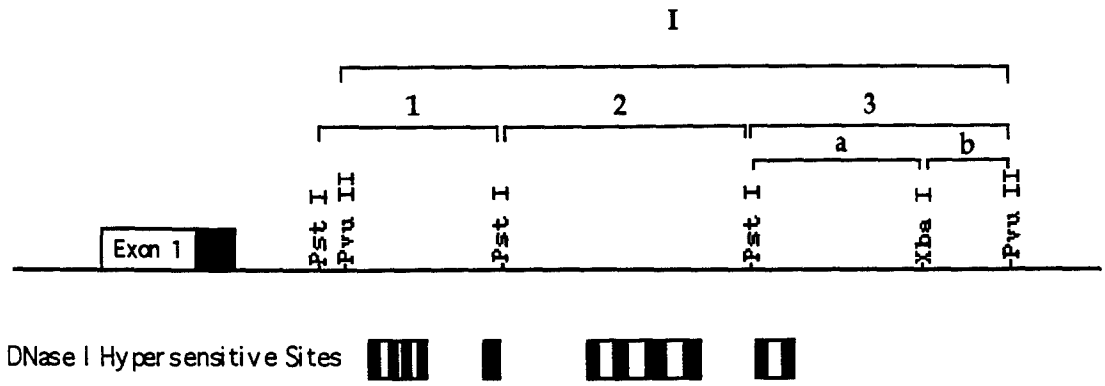


Figure 4. Comparison of CAT activity with the SV40 promoter-EGFR intron constructs in the forward orientation only in the MCF-7, T47D, HS578T and BT549 cell lines.



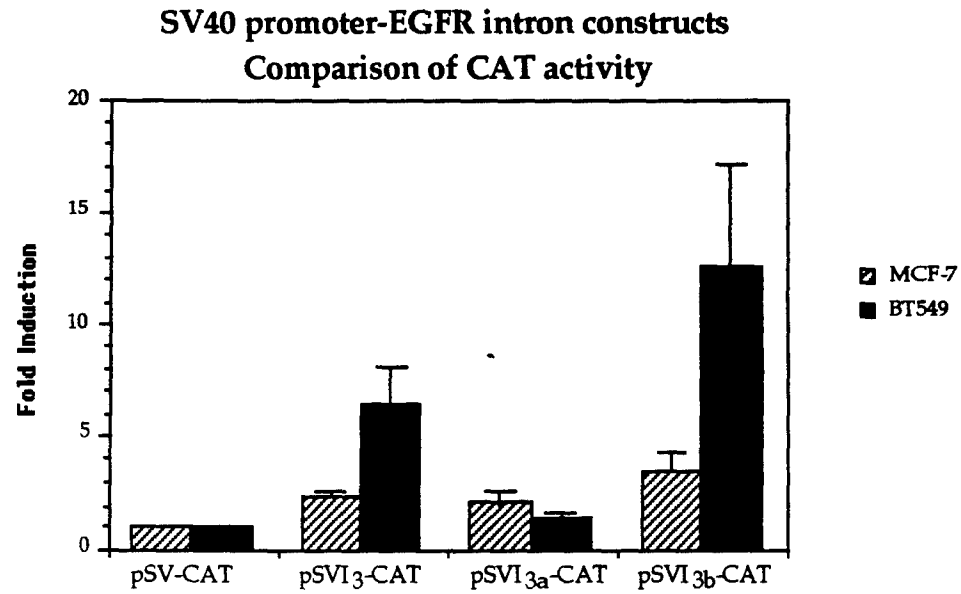
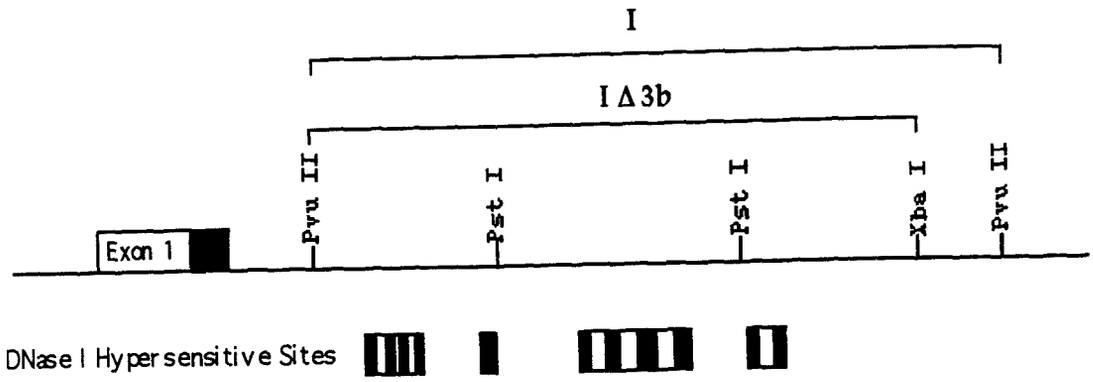


Figure 5. CAT activity seen in the MCF-7 and BT549 cell lines when region 3 is subdivided into 500bp and 140bp fragments and then subcloned into the pSV-CAT construct.



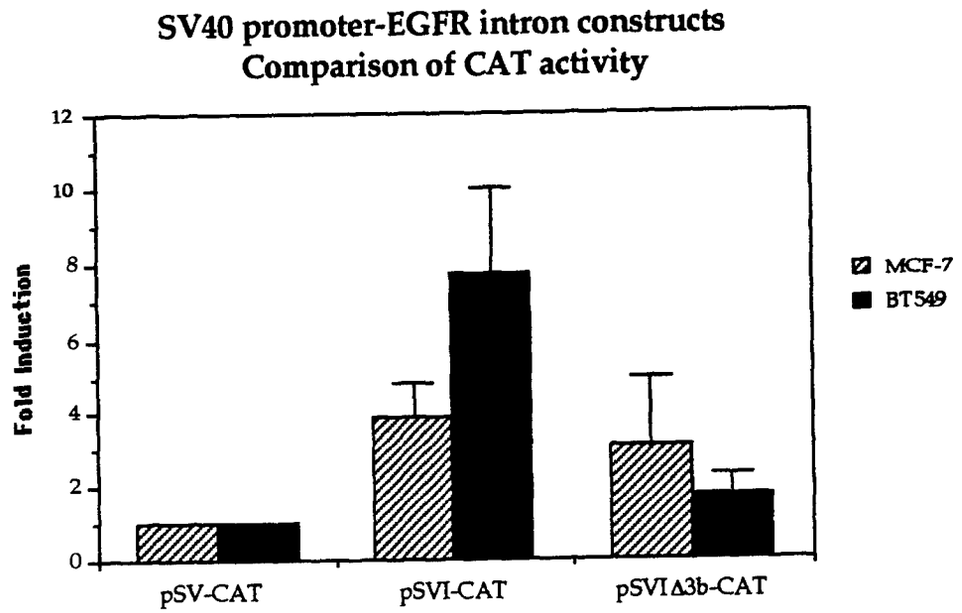


Figure 6. Comparison of CAT activity seen in the MCF-7 and BT549 cells when the entire 2.1kb intron region drives CAT expression as opposed to the 2.1kb region without the enhancer region 3b.

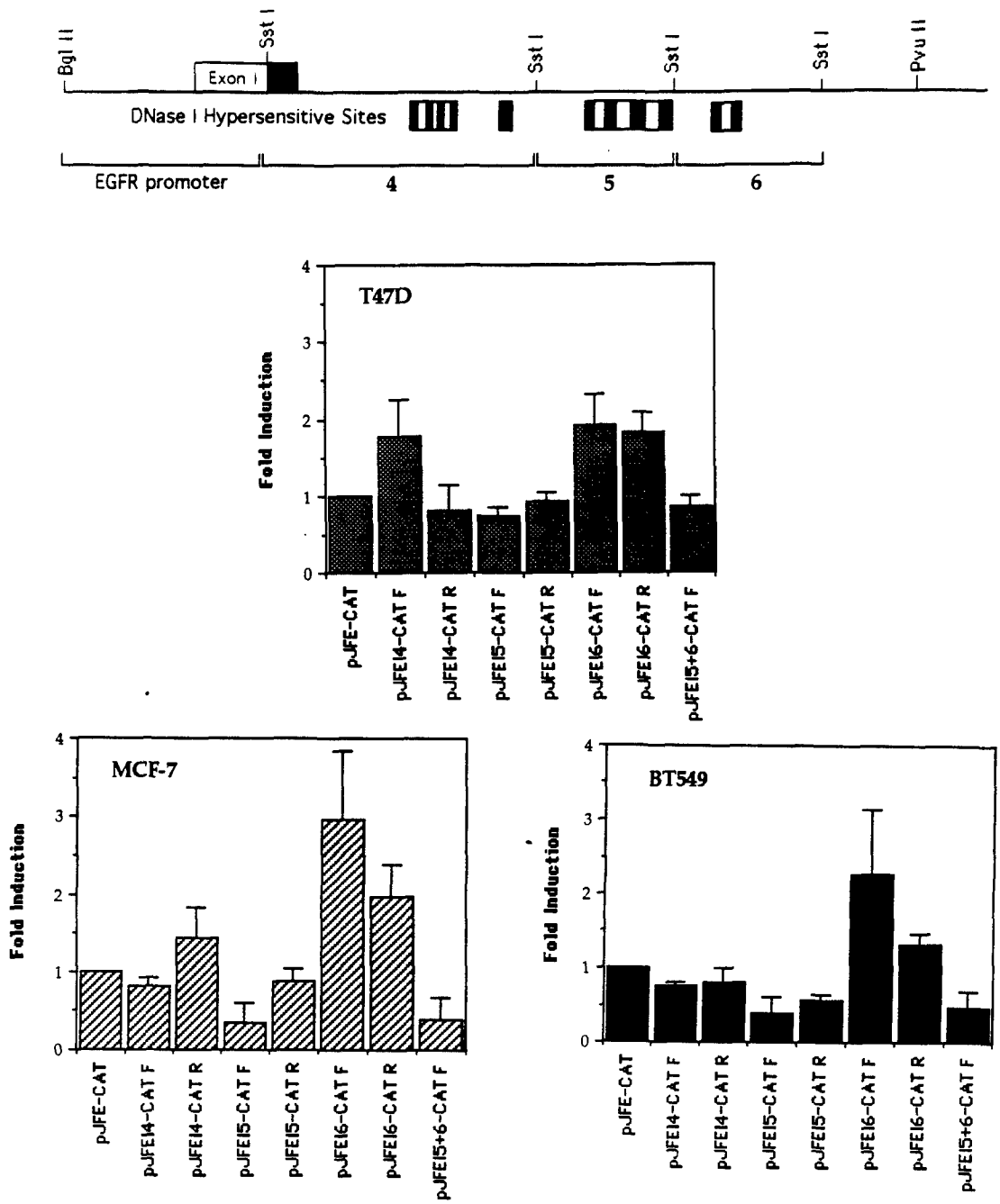
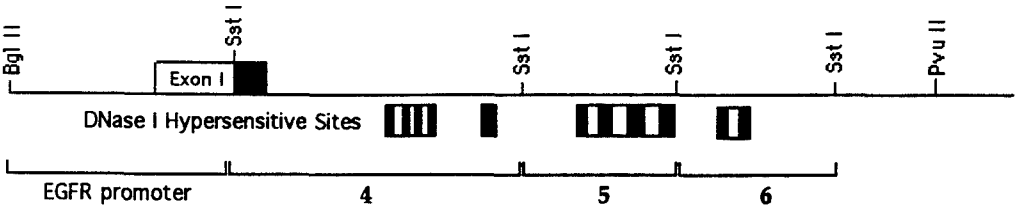


Figure 7. Relative CAT activity of the EGFR promoter-intron series of constructs in both the forward and reverse orientations in the T47D, MCF-7 and BT549 breast cancer cell lines.



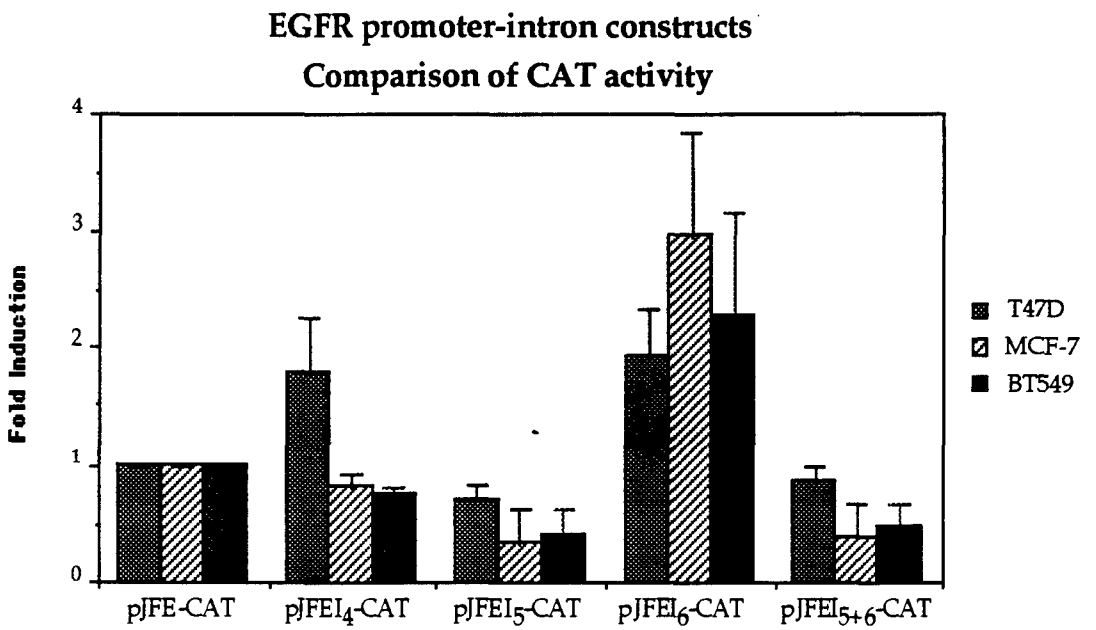


Figure 8. Comparison of CAT activity with the EGFR promoter-intron constructs in the forward orientation only in the T47D, MCF-7, and BT549 cell lines.

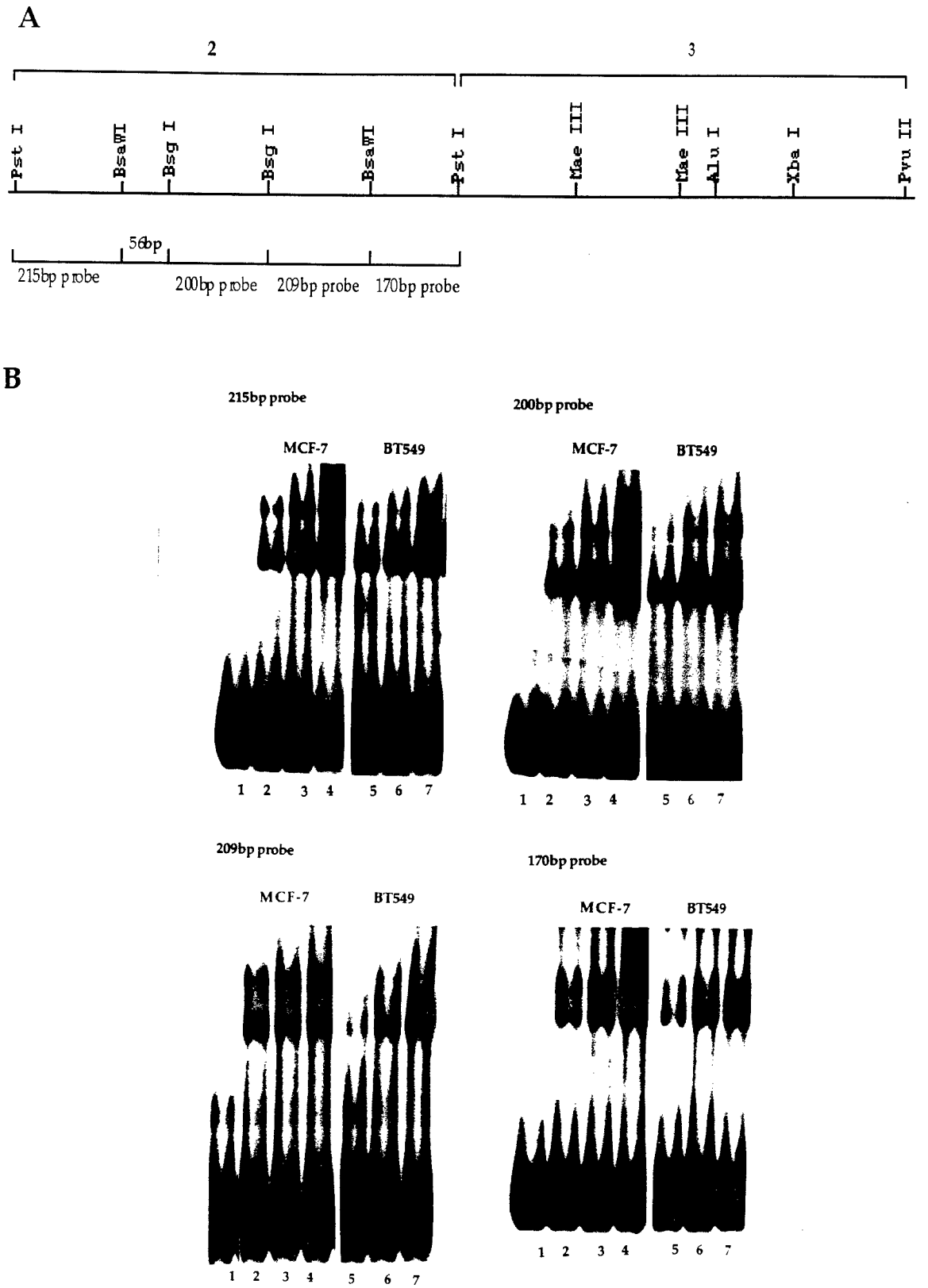


Figure 9. Representative gel shifts for region 2. (A) Schematic drawing of the regions used for probes. (B) All of the shifts are as follows: Lane 1 is probe alone; lanes 2-4 are probe incubated with MCF-7 nuclear extract and lanes 5-7 are probe incubated with BT549 nuclear extract; lanes 2 and 5 are 1mg of extract; lanes 3 and 6 are 2.5mg of extract; lanes 4 and 7 are 5mg of extract.

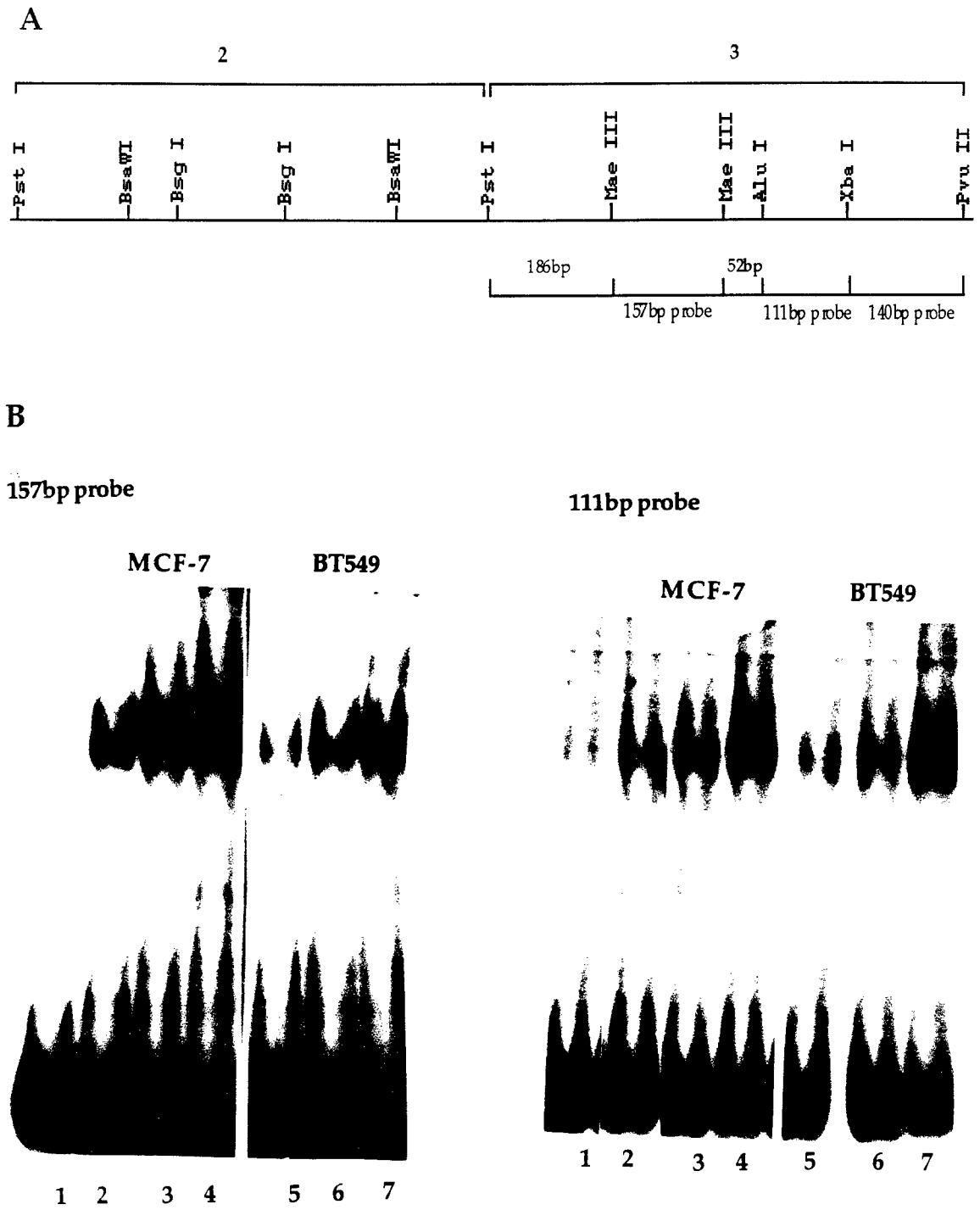
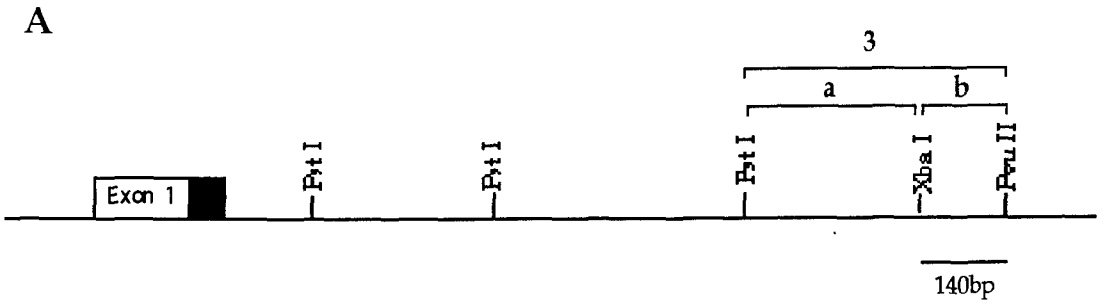


Figure 10. Representative gel shifts for region 3. (A) Schematic drawing of the regions used for probes. (B) Both shifts are as follows: Lane 1 is probe alone; lanes 2-4 are probe incubated with MCF-7 nuclear extract and lanes 5-7 are probe incubated with BT549 nuclear extract; lanes 2 and 5 are 1mg of extract; lanes 3 and 6 are 2.5mg of extract; lanes 4 and 7 are 5mg of extract.



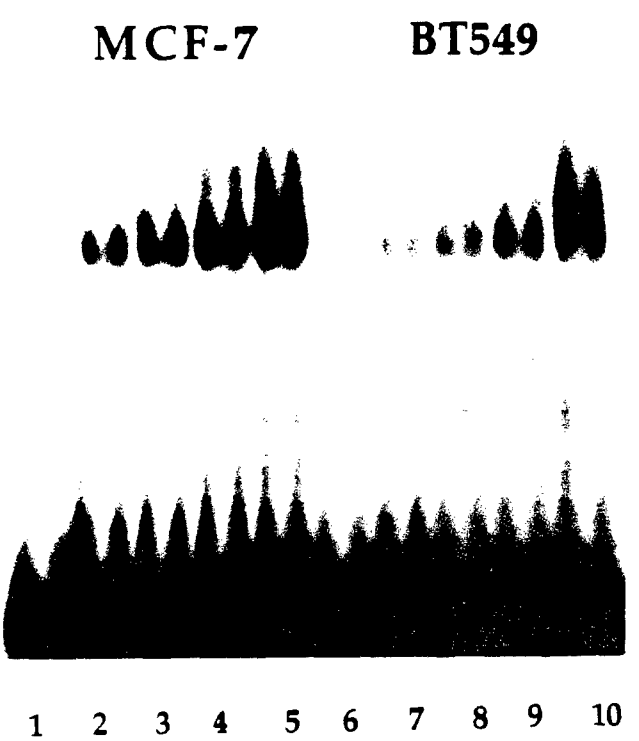
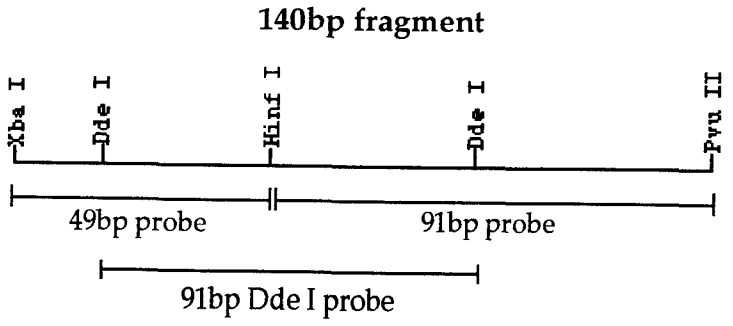


Figure 11. Gel shift of region 3b or the 140bp XbaI-PvuII fragment that showed enhancer activity with the SV40 promoter in BT549 cells. (A) Schematic drawing of the region of the intron used to perform the shift. (B) Gel shift of the 140bp probe. Lanes 1 and 6 are probe alone; lanes 2-5 are probe incubated with MCF-7 nuclear extract and lanes 7-10 are probe incubated with BT549 nuclear extract; lanes 2 and 7 are 1mg of extract; lanes 3 and 8 are 2.5mg of extract; lanes 4 and 9 are 5mg of extract; lanes 5 and 10 are 10mg of extract.

A



B

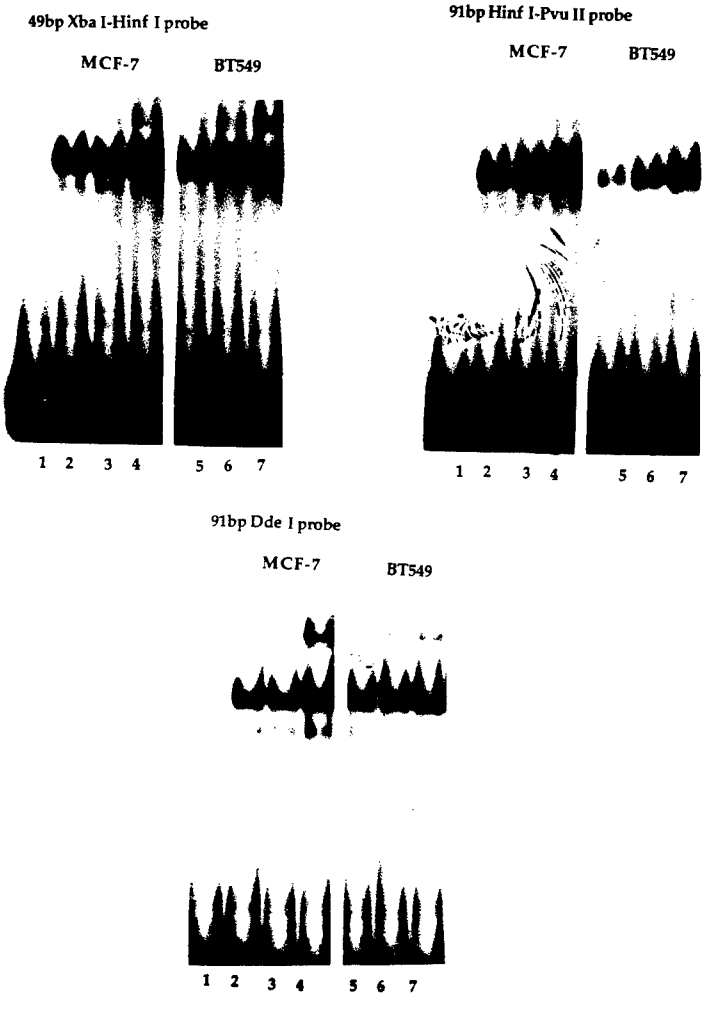
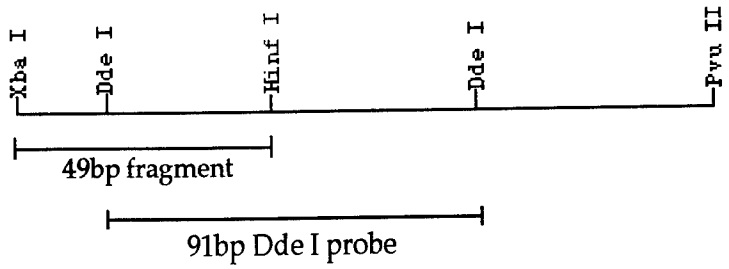


Figure 12. Gel shift of the subdivisions of the 140bp probe. (A) Schematic drawing of the regions used for probes. (B) All of the shifts are as follows: Lane 1 is probe alone; lanes 2-4 are probe incubated with MCF-7 nuclear extract and lanes 5-7 are probe incubated with BT549 nuclear extract; lanes 2 and 5 are 1mg of extract; lanes 3 and 6 are 2.5mg of extract; lanes 4 and 7 are 5mg of extract.

A



B

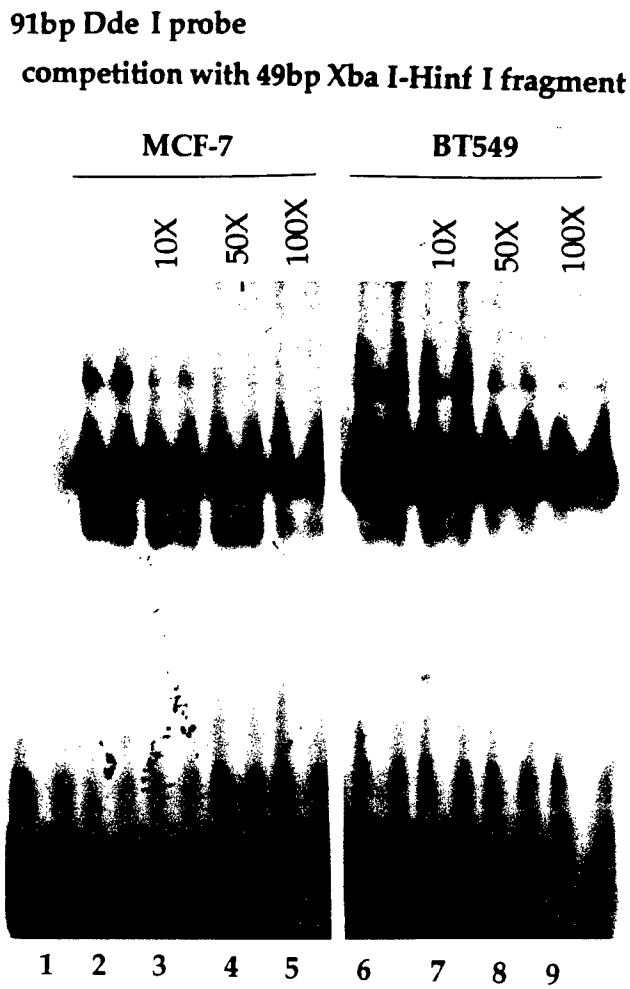


Figure 13. Competition of the 91bp DdeI probe with the 49bp fragment. (A) Schematic drawing of the regions used for the probe and the fragment used for competition. (B) Gel shift of the competition of the 91bp DdeI probe with the 49bp fragment. Lane 1 is probe alone; lanes 2-5 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 6-9 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 6 are 5mg of extract alone; lanes 3-5 and lanes 7-9 are competed with increasing amounts of the 49bp fragment.

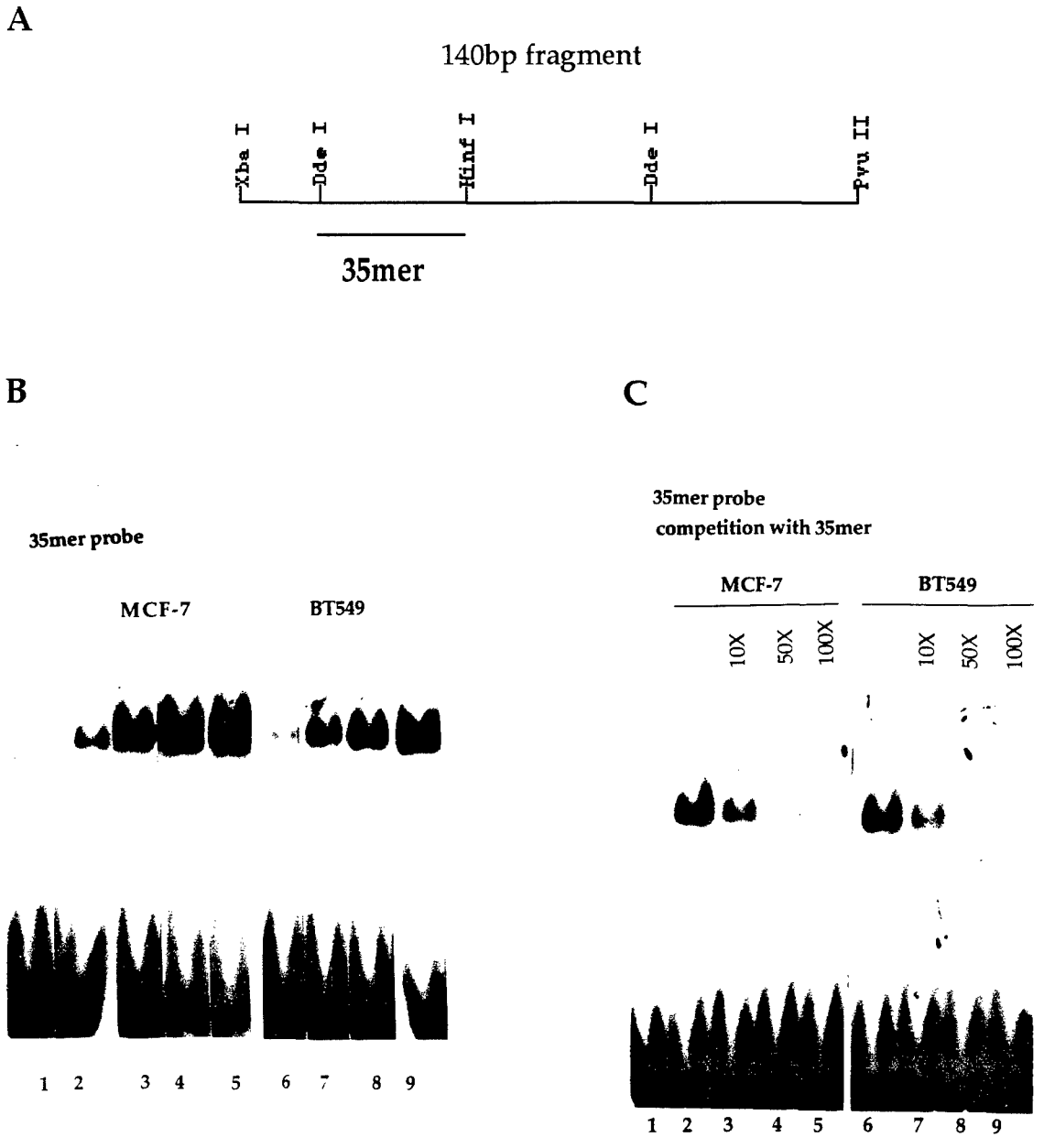


Figure 14. Gel shift and competition of the 35mer. (A) Schematic drawing of the 35mer within the 140bp fragment. (B) Gel shift with the 35mer as probe. Lane 1 is probe alone; lanes 2-5 are probe incubated with MCF-7 nuclear extract and lanes 6-9 are probe incubated with BT549 nuclear extract; lanes 2 and 6 are 1mg of extract; lanes 3 and 7 are 2.5mg of extract; lanes 4 and 8 are 5mg of extract; lanes 5 and 9 are 10mg of extract. (C) Competition of the 35mer with itself. Lane 1 is probe alone; lanes 2-5 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 6-9 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 6 are 5mg of extract alone; lanes 3-5 and lanes 7-9 are competed with increasing amounts of the 35mer.

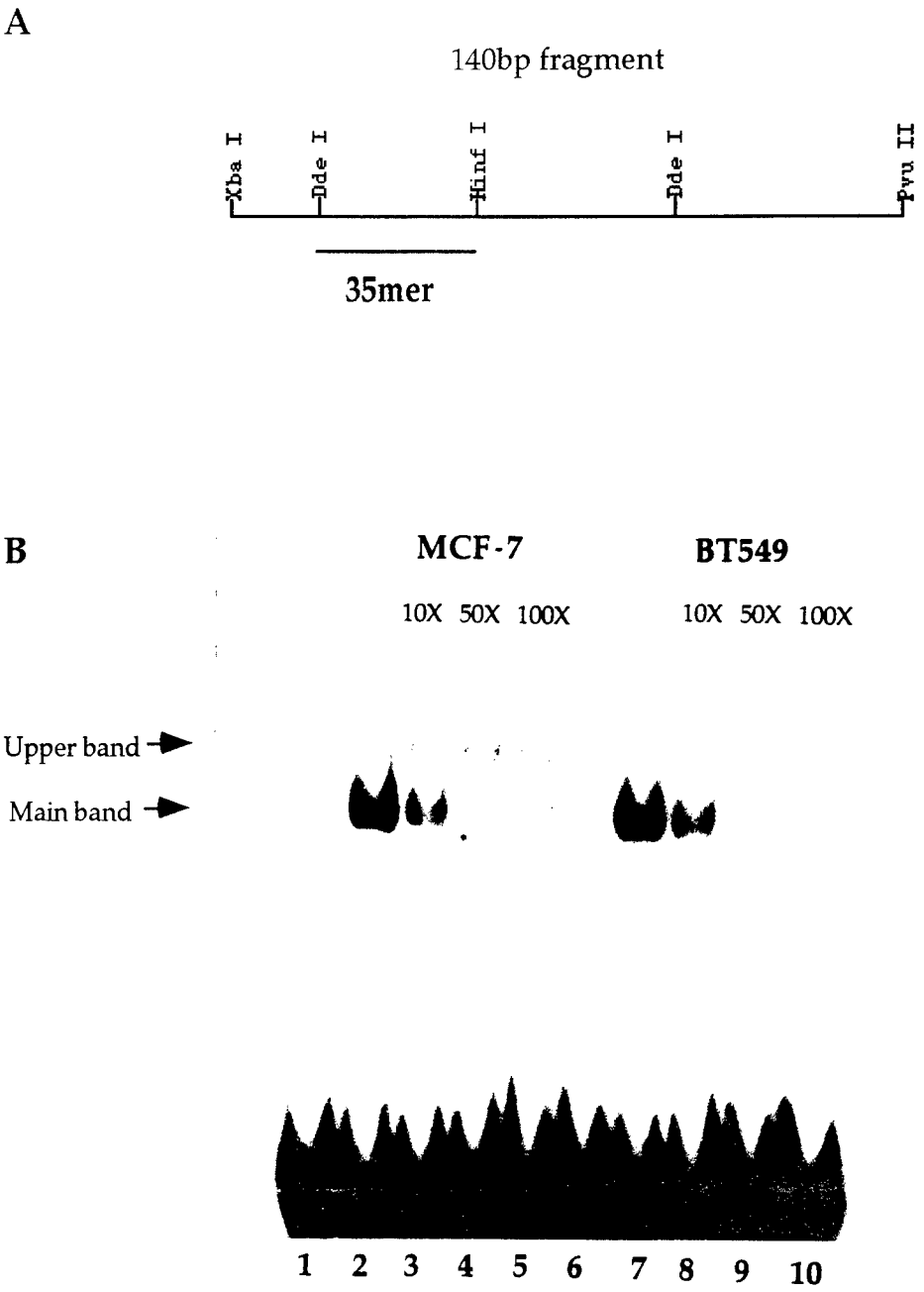


Figure 15. Competition of the 140bp XbaI-PvuII probe with 35mer. (A) Schematic drawing of the 35mer within the 140bp fragment. (B) Competition of the 140bp probe with the 35mer. Lanes 1 and 6 are probe alone; lanes 2-5 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 7-10 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 7 are 5mg of extract alone; lanes 3-5 and lanes 8-10 are competed with increasing amounts of the 35mer.

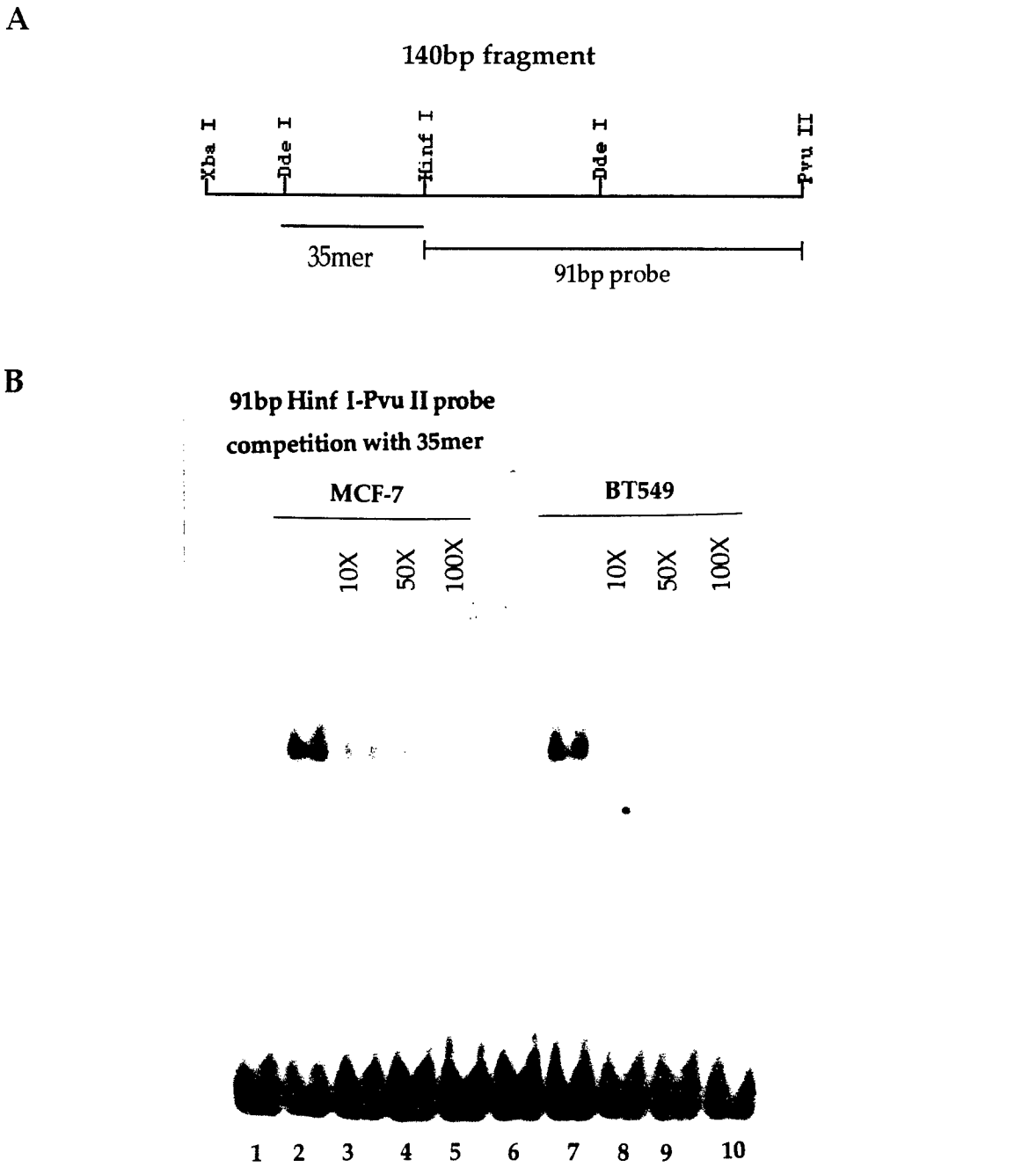


Figure 16. Competition of the 91bp Hinf I-PvuII probe with the 35mer. (A) Schematic drawing of the positioning of the 35mer and the 91bp Hinf I-PvuII fragment. (B) Competition of the 91bp Hinf I-PvuII probe with the 35mer. Lanes 1 and 6 are probe alone; lanes 2-5 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 7-10 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 7 are 5mg of extract alone; lanes 3-5 and lanes 8-10 are competed with increasing amounts of the 35mer.

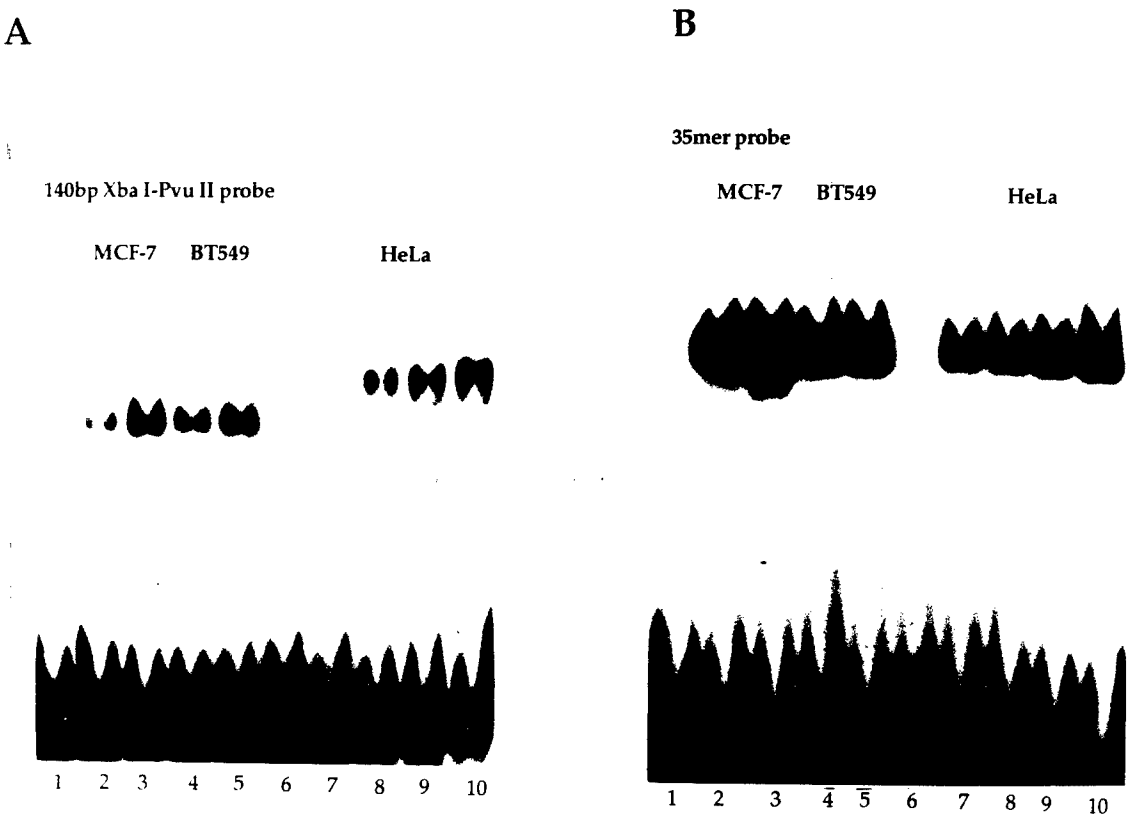
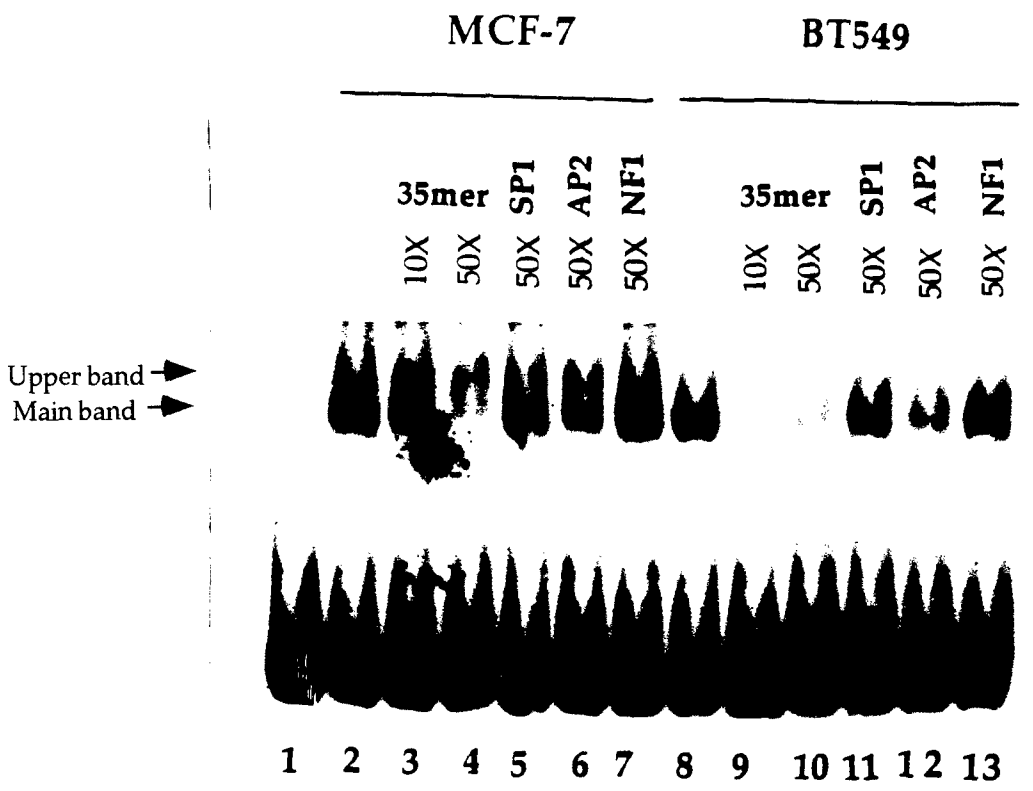


Figure 17. Comparison of gel shifts using extracts from HeLa cells and breast cancer cells. (A) Gel shift with 140bp XbaI-PvuII fragment as probe. Lanes 1 and 6 are probe alone; lanes 2 and 3 are probe incubated with 2.5mg and 5mg of MCF-7 nuclear extract respectively and lanes 4 and 5 are probe incubated with 2.5mg and 5mg of BT549 nuclear extract respectively; lanes 7-10 are incubated with HeLa nuclear extract at concentrations of 1mg, 2.5mg, 5mg and 10mg respectively. (B) Gel shift with 35mer as probe. Lanes are the same as in (A).

A



B

35mer	5'-CTTAGAGGTTATGACTGCCAAGACACCATTTCATG-3'
SP1	5'-GATCGATCGGGGCGGGGCGATC-3'
AP-2	5'-GATCGAACTGACCGCCCGCGGCCCGT-3'
NF-1	5'-ATTTTGGCTTTGAAGCCAATATG-3'

Figure 18. Competition of the 140bp XbaI-PvuII probe with the 35mer and oligonucleotides of common consensus cis elements. (A) Gel shift. Lane 1 is probe alone; lanes 2-7 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 8-13 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 8 are 5mg of extract alone; lanes 3-7 and lanes 9-13 are competed with the designated oligonucleotides. (B) The sequence of the oligonucleotides used for competition with the underlined bases as the consensus cis elements.

A

35mer
5'-CTTAGAGGTTATGACTGCCAAGACACCATTTCATG-3'

Mutant 35mer
5'-CTTAGAGGTTCATGTAGCCAAGACACCATTTCATG-3'

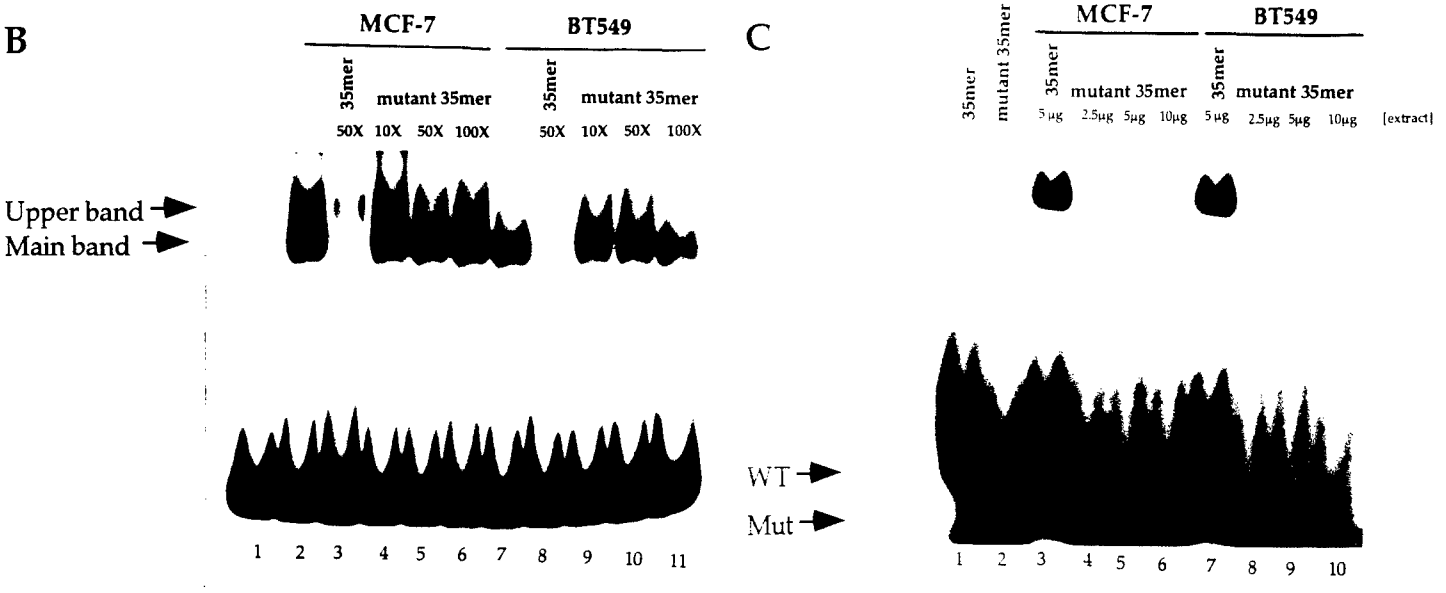


Figure 19. Competition and binding with the mutant 35mer. (A) Comparison of the sequences of the 35mer and the mutant 35mer. The underlined bases are the candidate cis element and the scrambled bases repectively. (B) Competition of the 140bp probe with the 35mer and the mutant 35mer. Lane 1 is probe alone; lanes 2-6 are probe incubated with 5mg of MCF-7 nuclear extract and lanes 7-11 are probe incubated with 5mg of BT549 nuclear extract; lanes 2 and 7 are 5mg of extract alone; lanes 3 and 8 are competed with 50X 35mer. Lanes 4-6 and lanes 9-11 are competed with increasing amounts of the mutant 35mer. (C) Gel shift with the 35mer and the mutant 35mer as probes. Lane 1 is the 35mer probe alone and lane 2 is the mutant 35mer probe alone; lanes 3-7 are probe incubated with MCF-7 nuclear extract and lanes 8-12 are probe incubated with BT549 nuclear extract; lanes 3 and 8 are 5mg of extract incubated with the 35mer probe; lanes 4 and 9 are 1mg of extract; lanes 5 and 10 are 2.5mg of extract; lanes 6 and 11 are 5mg of extract; lanes 7 and 12 are 10mg of extract incubated with the mutant 35mer probe.

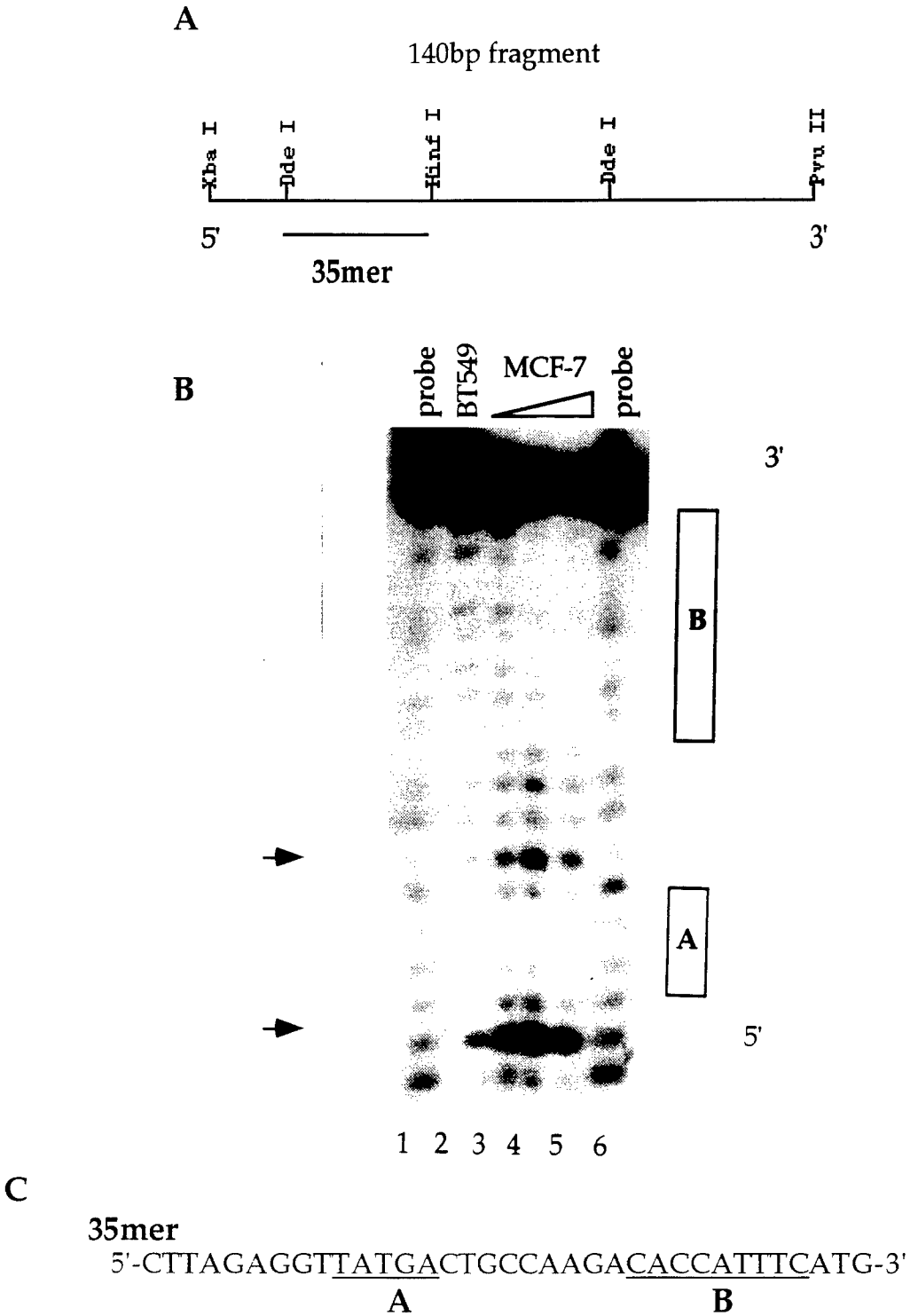


Figure 20. DNase I protection assay of the 35bp DdeI-HinfI fragment. (A) Schematic drawing of the 35bp fragment that was used as a probe. (B) Protection of the 35bp fragment that has been labeled at the 5' end. Lanes 1 and 6 are probe alone; lane 2 is probe incubated with 15mg of BT549 nuclear extract; lanes 3-5 are probe incubated with increasing amounts of MCF-7 nuclear extract, 20, 40, and 60mg respectively. Rectangles denote regions of protection. (C) The sequence of the 35mer with the underlined sequences denoting regions of protection.

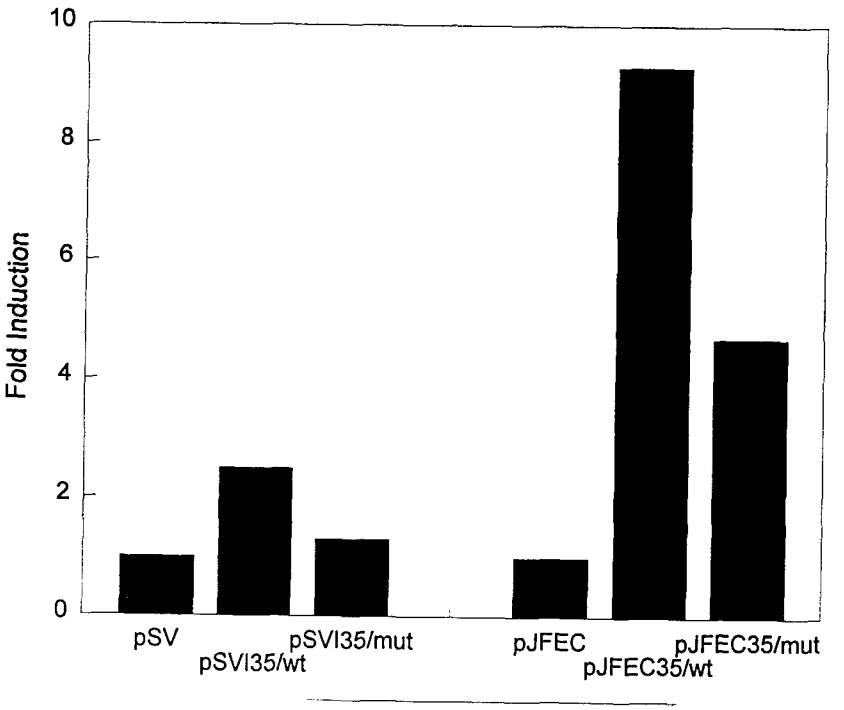


Figure 21. Relative CAT activity in BT549 cells of the wild type 35mer and the mutant 35mer subcloned into the SV40 promoter-CAT construct and the EGFR promoter-CAT construct.

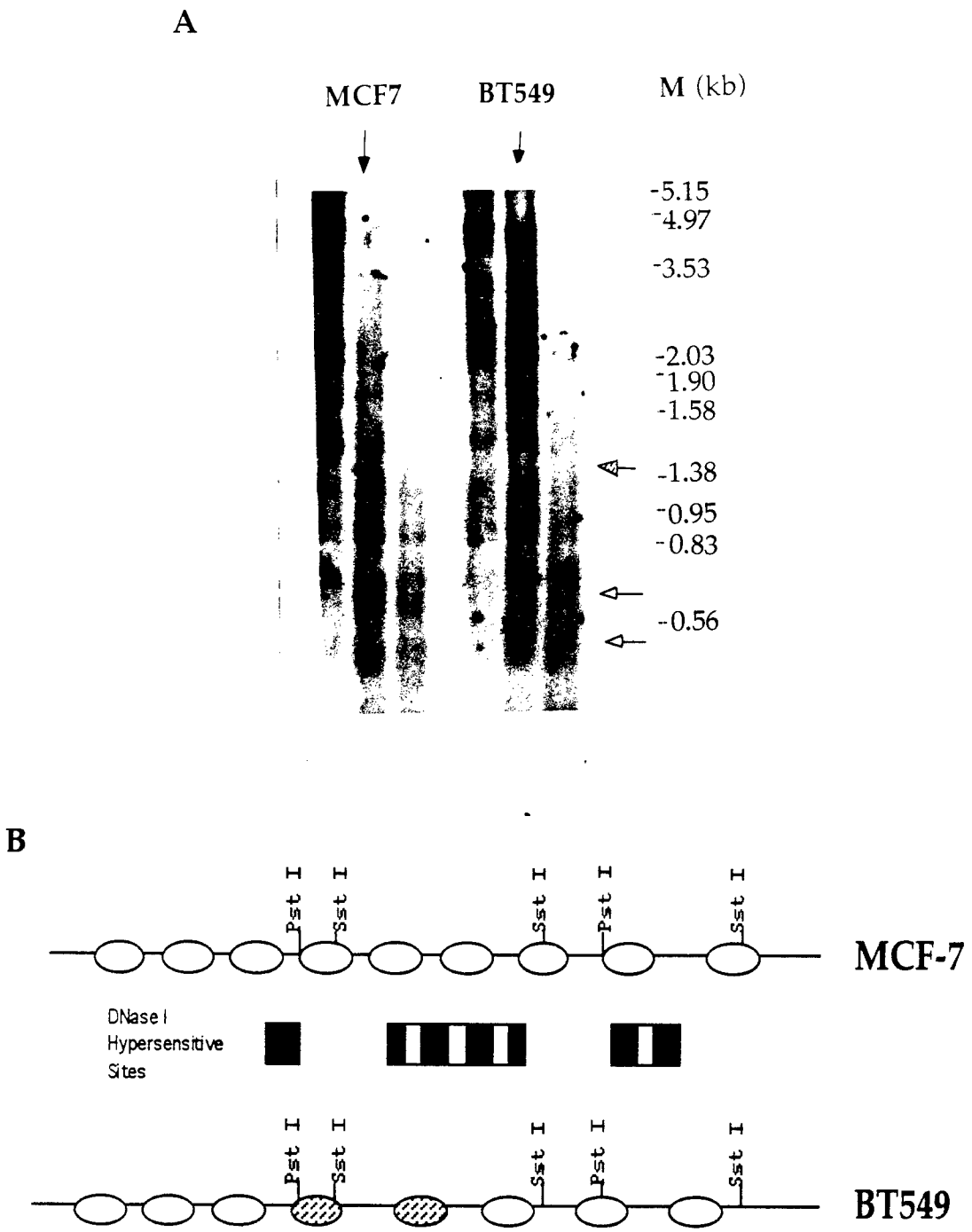


Figure 22. Nucleosomal phasing of the first intron of EGFR. The micrococcal nuclease assay was performed as described in Materials and Methods. (A) Southern blot of the micrococcal nuclease assay. The solid arrows denote the lanes for comparison and the shaded arrow depicts the region of nucleosome disruption within the phasing of the EGFR gene in BT549 cells. The open arrows indicate the region of the 80bp shift of nucleosomes. (B) This schematic model represents the mapping of the nucleosome phasing in the EGFR first intron in these low and high EGFR expressing cell lines. The shaded ovals are nucleosomes that have been disrupted in the BT549 cells. Also shown are the aforementioned DNase I hypersensitive sites.

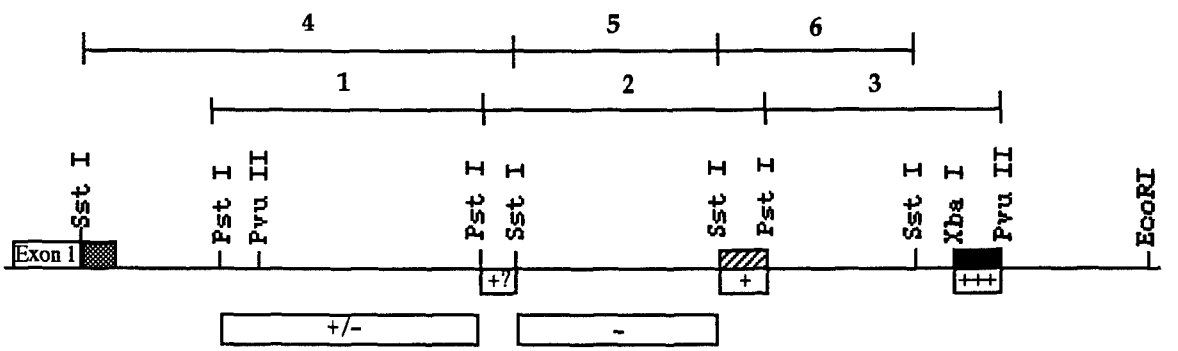


Figure 23. Schematic representation of the regulatory regions of the upstream first intron of the EGFR gene. The activities of these regions were predicted by combining the data from CAT assays performed with our two series of constructs. The activity of region 1 is dependent on the cell line used in the CAT assay. The striped box is a potential 136bp enhancer in breast cancer cells and HeLa cells as mentioned in the text. The black box is our 140bp enhancer in the BT549 cells.

BIBLIOGRAPHY

Presentations

Breast Cancer Risk Factors. Jane M. Murphey. Wilson College Homecoming Weekend, Breast Cancer Series October 1995

Poster Presentations

McInerney, J.M. and Chrysogelos, S.A. The role of first intron elements in epidermal growth factor receptor gene regulation in human breast cancer cells. The Era of Hope: The department of defense Breast Cancer Research Program Meeting, Washington, DC, October 31-November 4, 1997.

Wilson, M.A., McInerney, J.M., and Chrysogelos, S.A. "Transcriptional repression of the EGFR Gene in Hormone-Dependent Breast Cancer Cells", presented at the Gene Transcription and Therapeutic Intervention meeting sponsored by Cambridge Healthtech Institute, Loew's L'Enfant Plaza, Washington, DC, June 2-3, 1997.

Murphey, J.M., McGaffin K.R. and Chrysogelos, S.A. The nucleosomal organization of the first intron region is involved in EGFR gene regulation in breast cancer cells. Keystone Symposia. "Epigenetic Regulation of Transcription", April 1995.

Murphey, J. and Chrysogelos, S. Intron 1 elements are involved in EGFR gene regulation in breast cancer cells. Keystone Symposia "Breast and Prostate Cancer II", March 1994.

Publications

McInerney JM, Welch JN, Strand KJ, Chrysogelos S. Involvement of a 35bp intronic element in the regulation of EGFR in breast cancer cells. In preparation.

Murphey, J.M. and Chrysogelos, S.A. (1995) Direct measurement of transfection efficiency in transient transfection assays. Biotechniques, 18:967-969.

Murphey, J.M., McGaffin K.R. and Chrysogelos, S.A. (1995) The nucleosomal organization of the first intron region is involved in EGFR gene regulation in breast cancer cells. J. Cell. Biochem., 19D, 161.

Chrysogelos, S.A., Yarden, R.I., Lauber, A.H. and Murphey, J.M. (1994) Mechanisms of EGF receptor regulation in breast cancer cells. Breast Cancer Res. Treat., 31:227-36.

Murphey, J. and Chrysogelos, S. (1994) Intron 1 elements are involved in EGFR gene regulation in breast cancer cells. J. Cell. Biochem., 18D, 238.

PERSONNEL

Predoctoral students receiving pay:

Jane M. Murphey McInerney (1994-1996)

Melissa A. Wilson (1996-1997)

Mentor (no salary support):

Susan A. Chrysogelos, Ph.D.

Received 2/8/00

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y) 26 Jan 00

MEMORANDUM FOR Administrator, Defense Technical Information Center, ATTN: DTIC-OCA, 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports written for the following Awards.

DAMD17-86-C-6169	ADB116203
DAMD17-94-J-4056	ADB218947
DAMD17-94-J-4394	ADB220575
DAMD17-94-J-4358	ADB236080
DAMD17-94-J-4169	ADB236753
DAMD17-94-J-4049	ADB234453
DAMD17-94-J-4080	ADB218909
DAMD17-94-J-4080	ADB233428
DAMD17-94-J-4431	ADB220348
DAMD17-94-J-4335	ADB234557
DAMD17-94-J-4388	ADB218872
DAMD17-94-C-4081	ADB246577
DAMD17-94-J-4025	ADB238010
DAMD17-94-J-4080	ADB241898
MIPR 96MM6720	ADB240182
MIPR 96MM6720	ADB226818

Request the limited distribution statement for Accession Document Numbers be changed to "Approved for public release; distribution unlimited." These reports should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Virginia Miller at DSN 343-7327 or by email at virginia.miller@det.amedd.army.mil.

FOR THE COMMANDER:

PHYLIS M. RINEHART
Deputy Chief of Staff for
Information Management